Samuel Frühwirth, BSc

Chiral resolution via two-dimensional porous membranes: A preliminary computational screening of candidate materials

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Technical Physics

submitted to

Graz University of Technology

Supervisor

Ass.Prof. Mag.phil. Dipl.-Ing. Dr.phil. Dr.techn. Andreas Hauser

Institute of Experimental Physics

Graz, Februar 2017

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date

Signature

Acknowledgments

I wish to express my gratitude to my supervisor Andreas Hauser for his continued support, commitment and help. Further I offer my sincere appreciation for the opportunities provided as well as for the skills and knowledge Andreas shared with me. The completion of my thesis could not have been accomplished without discussions, ideas and motivation by my dear colleagues Johann and Ralf. I sincerely thank Professor Ernst for the possibility to write my thesis at his institute as well as for his support in the background. I also thank the Wirtschaftskammer Steiermark for supporting my thesis with a "Wirtschaft trifft Wissenschaft" grant.

Abstract

Recently, a new concept of chiral separation via functionalised, nanoporous graphene was developed by Hauser et al. (*Angew. Chem. Int. Ed.*, 2014, 53:9957-9960). In this work the possibility to apply this concept on pharmaceutically relevant systems is discussed together with the necessary computational techniques. Pure ab initio computations are limited to systems of a few hundred atoms, thus other simulation techniques are required. One possible choice are molecular mechanics methods. Therefore, the capability of reproducing the available ab initio results for the pore propagation is tested for the two standard force fields CGenFF and GAFF, which both are designed for the description of small drug molecules. Further tests include the formation of diastereomeric complexes build from various chiral compounds, studied with DFT and molecular mechanics methods, such as CGenFF, GAFF and MMFF94, with the aim to determine if force fields can describe the resulting geometries with the required accuracy. Another test comprises the propagation process itself. The transmission of small linear gas molecules through porous graphene and graphdiyne is calculated and the ab initio and molecular mechanics results are compared.

Kurzfassung

Unlängst wurde eine neues Konzept zur Trennung chiraler Moleküle mithilfe von funktionalisiertem, nanoporösem Graphen von Hauser et al. (*Angew. Chem. Int. Ed.*, 2014, 53:9957-9960) vorgestellt. In dieser Arbeit wird die Möglichkeit, dieses Konzept auf pharmazeutisch relevante Systeme anzuwenden, zusammen mit den dafür benötigten Berechnungsmethoden, untersucht. Da reine ab-initio Berechnungen nur auf Systeme von wenigen hundert Atomen angewendet werden können, müssen andere Simulationstechniken in Betracht gezogen werden. Insbesondere bieten sich hier molekularmechanische Methoden an. Zwei Standardkraftfelder, CGenFF und GAFF, die eigens für die Simulation von kleinen Arzneimittelmolekülen entwickelt wurden, werden auf ihre Übereinstimmung mit ab-initio Resultaten hin überprüft. Weitere Tests beinhalten die Bildung von diastereomerischen Komplexen, welche aus diversen chiralen Molekülen zusammengesetzt sind. Diese Komplexe werden mittels DFT und Molekularmechanikmethoden untersucht, wobei die Kraftfelder CGenFF, GAFF und MMFF94 zum Einsatz kommen. Zweck dieser Untersuchung ist die Überprüfung, ob Kraftfelder die optimierten Geomterien mit der notwendigen Genauigkeit wiedergeben können. Ein weiterer Test beinhaltet schließich den Porentransit, welcher die Transmission von kleinen linearen Gasmolekülen durch poröses Graphen und Graphdiyne umfasst. Die erhaltenen ab-initio Resultate werden erneut mit Kraftfeldberechnungen verglichen.

Das Leben ist wert, gelebt zu werden, sagt die Kunst, die schönste Verführerin; das Leben ist wert, erkannt zu werden, sagt die Wissenschaft.

Friedrich Nietzsche

Contents

List of Figures

List of Tables

List of Abbreviations

BSSE basis set superposition error

CGenFF CHARMM General Force Field

DFT density functional theory

FF force field

HMOF homochiral metal-organic framework

GAFF Generalized AMBER Force Field

GGA generalised gradient approximation

HF Hartree-Fock

HPLC high-performance liquid chromatography

KS Kohn-Sham

LDA local density approximation

LJ Lennard-Jones

LSDA local spin density approximation

MD molecular dynamics

MEP minimum energy path

MM molecular mechanics

MMFF94 Merck Molecular Force Field 94

MO molecular orbital

NEB nudged elastic band

QC quantum chemical

QMMM quantum mechanics/molecular mechanics

TS transition state

vdW van der Waals

1 Introduction

Enantiomers, the mirror images of a drug molecule, affect biological systems in different ways, although the majority of their chemical and physical properties is identical. For example, it is possible that one enantiomer acts antibiotic while the other does not have any effect on the human body. One form could help curing diseases while the other could be toxic. At least since the Contergan tragedy in the 1960s in Germany the pharmaceutical industry is fully aware of the importance of chiral resolution. The active pharmaceutical ingredient of Contergan is thalidomide, which occurs in two forms: (+)-thalidomide, which is sedative, and (−)-thalidomide, which is teratogenic and leads to deformation of the fetus when taken during a specific stage of the pregnancy. Therefore, it is an essential requirement for the chemical industry to separate and test all enantiomers of a bioactive substance [1, 2]. Around 40% of drugs are chiral compounds and 12% are administered as pure single enantiomers. This situation is gradually changing as a number of companies have now started to move towards producing enantiomerically pure forms of established drugs [3]. In most cases, a direct enantioselective synthesis is not feasible and intermediate forms or final products need to be extracted from racemic mixtures. This leads to new challenges with respect to separation methods. Common techniques comprise asymmetric synthesis, HPLC, crystallisation or kinetic resolution. Asymmetric synthesis may be the best choice for many cases as it yields 100% of one enantiomer and thus is very effective. However, this is not possible for all chiral drugs. Other methods are either not cost-effective or not universally applicable. As a consequence the development of new methods is inevitable. For that purpose a new method, which is based on the three-point model [4, 5], was developed by Hauser et al [6]. The principle of this method is to separate enantiomers with a porous, two-dimensional, single-atom thick membrane, which was functionalised with a chiral compound on the pore rims. These chiral compounds act as gatekeepers. They form diastereomeric complexes with the dissolved enantiomers, which then have a different diameter and thus either fit through the pore or not, leading to a separation effect. The importance of matching the pore size with

the size of the chiral molecule for enantioselective adsorption has recently been studied for homochiral metal-organic frameworks (HMOFs) [7]. HMOFs are nanoporous materials with metallic compounds and chiral ligands as linkers. The successful synthesis of more than 30 HMOFs has been reported [8], making them an interesting and probably more realistic class of materials for the given task than a single-layered membrane. However, in this thesis we focus on nanoporous sheets of graphene as well-studied benchmark system. Note that pristine graphene sheets [9] are impermeable for all types of gases. Hence, the necessary pores have to be generated either by post-synthesis methods or by direct synthesis of the desired structure [10, 11, 12, 13, 14]. Common techniques are hole drilling with electron beams, UV-induced oxidative etching [15, 16, 17], ion bombardment [18, 19], or self-assembling [20, 21, 22] with suitable precursors. Recently, a membrane with a desirable pore diameter for drug molecules has been synthesised [23]. In order to extend the concept of chiral separation to treat pharmaceutically relevant systems of a few hundred up to thousands of atoms, suitable simulation techniques have to be applied. Unfortunately the most accurate class of methods, quantum chemistry calculations, can not be applied to systems of that size. Therefore, classical molecular mechanics or mixed quantum mechanics/molecular mechanics approaches are potential techniques, but have to be tested for their capability of reproducing pure ab initio data within an acceptable error before applying them on a larger scale. It is the aim of this thesis to study the performance of common force fields in the given context, and to identify first molecular templates, which seem suitable for membrane-based chiral separation.

2 Theory and methods

2.1 Molecular mechanics

2.1.1 The idea

Molecular mechanics (MM) uses classical mechanics to model molecular systems, which is particularly useful for systems that are too large for quantum chemical (QC) methods ($N >> 10^2$) and/or have too long-lasting simulation times for ab initio molecular dynamics (MD) simulations. The term "classical mechanics" means that the system consists of particles, typically atoms or groups of atoms, with their interactions described by empirically adjusted force fields. The main application of MM are MD simulations, which allow the computation of equilibrium and transport properties of a molecular system. The usual workflow in an MD simulation is to select a model system consisting of N particles and to solve Newton's equations of motion for the system until equilibrium is reached. Afterwards, the simulation is continued under the desired conditions to measure its properties in a statistical manner. Therefore, the observable needs to be expressed as a function of the position and momenta of the particles [24].

2.1.2 Force fields: basic structure and common forms

The potential energy surface in an MM simulation is calculated by applying force fields (FFs). Each FF uses a specific function to describe the potential energy of the system, consisting of bonded and non-bonded terms. The force acting on the particles is given by the negative gradient of this energy,

$$\vec{F} = -\nabla E. \tag{2.1}$$

Table 2.1: Overview of the areas of application for common force fields and force field families

Force field(s)	Molecules
AMBER	proteins, nucleic acids, carbohydrates, general
CFF91/93/95	general, polymers and metals
CHARMM	proteins, carbohydrates, lipids, general
CVFF	general
DREIDING	general
GROMOS	proteins, nucleic acids, carbohydrates, lipids, polymers
MM2/MM3/MM4	general, all elements
MMFF94	general
OPLS	proteins, nucleic acids
TraPPE	C,N,O compounds
TRIPOS	general
UFF	all elements

A typical choice for these functions are polynomials for bonds and angles, Fourier series for the dihedrals and Coulomb and Lennard-Jones (LJ) potentials for the non-bonded interactions. The parameters of a FF are derived from ab initio or experimental data towards application on specific molecular systems, e.g. proteins, carbohydrates, polymers or metals. Hence, the choice of a suitable FF for the actual system is crucial. Widely spread FFs in the area of bio-chemistry are the AMBER and CHARMM FFs. In this study, one widely applicable FF and two FFs, developed especially for the simulation of small organic molecules, are used for our molecular systems, which are introduced in sections 2.1.3, 2.1.4 and 2.1.5.

2.1.3 CHARMM General Force Field

The CHARMM General Force Field (CGenFF) [25,26] is an all-atom force field specifically developed to simulate small drug-like molecules. All-atom means that every atom is treated as a single particle in the MM calculation, in contrast to united-atom FFs in which whole groups of atoms are treated as one particle. CGenFF offers a vast variety of different atom types and an online tool to derive missing parameters by analogy and net charges of the

atoms [27, 28]. The corresponding energy function for CGenFF is represented in Eq. 2.2.

$$\begin{aligned} \mathrm{E_{pot}} = & \underbrace{\sum_{\text{bonds}} K_b \Delta b^2}_{\mathrm{E_{bond}}} + \underbrace{\sum_{\text{angles}} K_\theta \Delta\theta^2}_{\mathrm{E_{angle}}} + \underbrace{\sum_{\text{dihedrals}} \frac{K_\phi}{2}(1 + cos(n\phi - \delta))}_{\mathrm{E_{dihedral}}} + \\ & \underbrace{\sum_{\text{impropers}} K_{\chi,ijk:l}\chi^2_{ijk:l}}_{\mathrm{E_{improper}}} + \underbrace{\sum_{\text{angles}} K_{\mathrm{UB}}(r_{13} - r_{0,13})^2}_{\mathrm{E_{Urey-Bradley}}} + \underbrace{\sum_{i<j} \frac{q_i q_j}{4\pi D r_{ij}}}_{\mathrm{E_{electrostatic}}} + \\ & \underbrace{\sum_{i<j} \epsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right]}_{\mathrm{E_{vdW}}}, \end{aligned} \tag{2.2}$$

$$\text{with} \quad \Delta b = b - b_0, \qquad \Delta\theta = \theta - \theta_0.$$

The first term describes the bond between two atoms with a harmonic oscillator approach. K_b denotes the force constant, b_0 is the equilibrium distance. The second term characterises the angular energy contribution with the angle being formed by three consecutively bonded atoms. K_θ is the force constant, θ_0 is the equilibrium angle. The third term describes the dihedral energy contribution. The dihedral is the angle between the two planes i-j-k and j-k-l formed by four atoms i,j,k,l, which are bonded as i-j, j-k, k-l. K_ϕ is the force constant, n the rotational symmetry, and δ the phase angle. The fourth term is the improper or out-of-plane energy. Considering four atoms i,j,k,l that are bonded i-l,j-l and k-l the out-of-plane angle is defined as the angle between the plane i,j,k and plane j,k,l. K_ϕ is the force constant, ϕ_0 is the the equilibrium angle. The fifth and last intramolecular term of the energy function is the Urey-Bradley potential. It is added as a quasi-cross-term for combined angle stretching and bending. A virtual bond between atoms i,k in an angle formed by atoms i,j,k is introduced. $r_{1,3;0}$ is the equilibrium distance and K_{UB} is the corresponding force constant. The second-last term is the electrostatic energy, with q_i, q_j as the net charges of two interacting atoms and D as the dielectric constant. The last part is the van der Waals (vdW) interaction, modelled by a standard 12-6 Lennard-Jones Potential. The well depth of the potential is defined as ϵ_{ij}, the equilibrium distance as $R_{min,ij}$. The Lorentz-Berthelot mixing rules [29] are applied for determining the cross-interaction parameters of the Lennard-Jones potential. No vdW or Coulomb potentials are applied on atoms directly bonded (1-2 interaction) or only separated by one atom (1-3 interaction), while they are fully taken into account for atoms with 2 or more bonded atoms

in between (1-4 interactions).

2.1.4 Generalized AMBER Force Field

The Generalized AMBER Force Field (GAFF) [30] is part of the AMBER FF family. It was designed for the same purpose as the CGenFF, namely to extend the applicability of already existing AMBER FFs for pharmaceutically relevant drug molecules. It has, compared to other FFs, a small number of atom types, but all possible parameters are included, hence none have to be derived by analogy or from equivalence lists. This easy applicability comes at the cost of decreased accuracy. The free software tool Antechamber [31] assists with the atom type assignment and partial charge calculation from either ab initio electrostatic potentials or with semi-empirical AM1-bcc computations. The energy function of GAFF has the form

$$\begin{aligned} \mathrm{E_{pot}} = & \underbrace{\sum_{\text{bonds}} K_b \Delta b^2}_{\mathrm{E_{bond}}} + \underbrace{\sum_{\text{angles}} K_\theta \Delta\theta^2}_{\mathrm{E_{angle}}} + \underbrace{\sum_{\text{dihedrals}} \frac{K_\phi}{2}(1 + cos(n\phi - \delta))}_{\mathrm{E_{dihedral}}} + \\ & \underbrace{\sum_{i<j} \frac{q_i q_j}{D r_{ij}}}_{\mathrm{E_{electrostatic}}} + \underbrace{\sum_{i<j} \epsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right]}_{\mathrm{E_{vdW}}}, \\ & \text{with} \quad \Delta b = b - b_0, \qquad \Delta\theta = \theta - \theta_0. \end{aligned} \tag{2.3}$$

Here, b_0 and Θ_0 are the equilibrium parameters for the bonds and angles, Ks are the force constants, n is the rotational symmetry and δ is the phase angle for the torsional angle parameters. $R_{min,ij}$ and ϵ_{ij} characterise the Lennard-Jones potential, with net charges q_i and q_j and the dielectric constant D for the electrostatic potential. The Lorentz-Berthelot mixing rules are applied for determining the cross-interaction parameters of the Lennard-Jones potential. The 1-2 and 1-3 intermolecular interactions are included in the bonded parameters. The 1-4 interactions coulomb interactions are scaled by a factor of $\frac{5}{6}$ and the 1-4 LJ interactions are scaled by a factor of 0.5.

2.1.5 Merck Molecular Force Field 94

The MMFF94 [32] was developed for a wide range of organic molecules in order to study macromolecule - small molecule interactions. The core portion of MMFF94 was built on ab initio results; only for the validation experimental data was used. Also, the functional form of the energy differs:

$$
\begin{aligned}
\mathrm{E_{pot}} = \\
&\underbrace{\sum_{\text{bonds}} \frac{K_b}{2}\Delta b(1 - 2\Delta b + \frac{7}{3}\Delta b^2)}_{\mathrm{E_{bond}}} + \underbrace{\sum_{\text{angles}} \frac{K_\theta}{2}\Delta\theta^2\left(1 - 0.4\left(\Delta\theta\right)\right)}_{\mathrm{E_{angle}}} + \\
&\underbrace{\sum_{\text{angles}} \left(K_{b\theta,ijk}\Delta b_{ij} + K_{b\theta,kji}\Delta b_{kj}\right)\Delta\theta_{ijk}}_{\mathrm{E_{stretch-bend}}} + \underbrace{\sum_{\text{impropers}} \frac{K_{\chi,ijk:l}}{2}\chi^2_{ijk:l}}_{\mathrm{E_{improper}}} + \\
&\underbrace{\sum_{\text{dihedrals}} \frac{1}{2}\left(K_1\left(1 + cos(\phi)\right) + K_2\left(1 - cos(2\phi)\right) + K_3\left(1 + cos(3\phi)\right)\right)}_{\mathrm{E_{dihedral}}} + \\
&\underbrace{\sum_{i<j} \frac{q_i q_j}{D\left(r_{ij} + \delta\right)^n}}_{\mathrm{E_{elecrostatic}}} + \underbrace{\sum_{i<j} \epsilon_{ij}\left(\frac{1.07R_{0,ij}}{R_{ij} + 0.07R_{0,ij}}\right)^7\left(\frac{1.12R^7_{0,ij}}{R^7_{ij} + 0.12R^7_{0,ij}} - 2\right)}_{\mathrm{E_{vdW}}}, \\
&\text{with} \quad \Delta b = b - b_0, \qquad \Delta\theta = \theta - \theta_0
\end{aligned}
\tag{2.4}
$$

A polynomial of the order four is chosen for the bond energy. The angle bending is described by a 3rd order polynomial. Contrary to the other FFs, the MMFF94 contains a cross-term: the stretch-bending, which couples the bending of an angle i-j-k and the stretching of the bonds i-j and j-k. The out-of-plane bending energy is described by a quadratic potential and $\chi_{ijk:l}$ is the Wilson angle [33] between plane i-j-k and the bond j-l, with l being the central atom. The dihedrals are described by a Fourier series expansion. The MMFF94 employs a special "buffered 14-7" [34] form for the van der Waals interactions with own special combination rules. The electrostatic potential has a buffered coulombic form with the "buffering" constant $\delta = 0.05\,\text{Å}$. D is the dielectric constant and n is usually taken as 1. Another distinction from the other force fields is the calculation of the partial charges, which are constructed from initial atomic charges by adding contributions from the attached bonds. The 1-4 vdW parameters are not scaled differently, but the 1-4 electrostatic interactions are scaled by a factor of 0.75.

2.1.6 Integrating the equations of motion: The Verlet algorithm and its derivations

In order to perform an MD simulation it is necessary to integrate Newton's equations of motion in an efficient way. One of the simplest algorithms is the Verlet algorithm. It has several advantages:

- It is, analogue to Newton's equations of motion, time reversible.
- It conserves a pseudo-Hamiltonian that approaches the real Hamiltonian of the system in the limit of infinitely small steps and machine precision. This results in a slight long-term energy drift.
- It requires minimal memory, although memory is not a real limitation in MD simulations.

The derivation of the following formulas follows the book of Frenkel and Smit [24]. Starting with a Taylor expansion of the coordinate of a particle around time t,

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{F}(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{3!}\dddot{\mathbf{r}} + O(\Delta t^4), \tag{2.5}$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \frac{\mathbf{F}(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{3!}\dddot{\mathbf{r}} + O(\Delta t^4), \tag{2.6}$$

we sum over Eq. 2.5 and 2.6 to obtain

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \frac{\mathbf{F}(t)}{2m}\Delta t^2 + O(\Delta t^4). \tag{2.7}$$

The velocity and the third order term cancel out, leading to an error of the order of Δt^4, where Δt is the time-step in the MD simulation. The velocity is not needed for the Verlet algorithm, but it is crucial to determine the kinetic energy and therewith the temperature. The velocities can be obtained from the trajectory via

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t} + O(\Delta t^2). \tag{2.8}$$

The error of the velocities is of the order of Δt^2 and needs extra calculations. Another disadvantage is that at $t = 0$, $\mathbf{r}(-\Delta t)$ is required and the trajectory can be affected by the initial guess; hence other Verlet-like algorithms are often used.

The leap frog algorithm

The leap frog algorithm is derived from the Verlet algorithm by defining the velocities at half integer time steps and then obtaining the new positions based on the old positions and velocities,

$$\mathbf{v}(t + \Delta t/2) = \mathbf{v}(t - \Delta t/2) + \frac{\mathbf{F}(t)}{m}\Delta t. \tag{2.9}$$

The velocities are updated following this expression:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t + \Delta t/2)\Delta t. \tag{2.10}$$

The leap frog algorithm offers an increased accuracy for the velocities but does obviously not solve the initial step problem. Since velocities are not defined at the same time as the positions, the kinetic and potential energy are also not defined at the same time. Hence, one cannot directly compute the total energy of the system. Both problems are solved by the velocity Verlet algorithm.

The velocity Verlet algorithm

The velocity Verlet algorithm is the standard algorithm used in most MD program packages. It looks like a Taylor expansion for the coordinates in discrete timesteps:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{F}(t)}{2m}\Delta t^2. \tag{2.11}$$

The update of the velocities is expressed as

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{\mathbf{F}(t + \Delta t) + \mathbf{F}(t)}{2m}\Delta t. \tag{2.12}$$

As can be seen, the new velocities can only be computed after the new positions are calculated, and from these the new forces are calculated.

2.2 Quantum chemical methods

2.2.1 From the time-independent Schrödinger equation to the Hartree-Fock equations

The derivation of the Hartree-Fock equations follows the book of F. Jensen [35]. Starting with the Schrödinger equation,

$$\mathbf{H}\Psi = E\Psi, \tag{2.13}$$

the total non-relativistic Hamilton operator can be written as a sum of kinetic and potential energy operators of the nuclei and electrons,

$$\begin{aligned}
&\mathbf{H}_{\text{tot}} = \mathbf{T}_{\text{n}} + \mathbf{H}_{\text{e}}, \quad \text{with} \\
&\mathbf{H}_{\text{e}} = \mathbf{T}_{\text{e}} + \mathbf{V}_{\text{ne}} + \mathbf{V}_{\text{ee}} + \mathbf{V}_{\text{nn}} \\
&\mathbf{T}_{\text{n}} = -\sum_{A} \frac{1}{2M_A}\nabla_A^2 =, \quad \mathbf{T}_{\text{e}} = -\sum_{i} \frac{1}{2M_i}\nabla_i^2 \\
&\mathbf{V}_{\text{ee}} = \sum_{i>j} \frac{1}{r_{ij}}, \quad \mathbf{V}_{\text{ne}} = -\sum_{i,A} \frac{Z_A}{r_{iA}}, \quad \mathbf{V}_{\text{nn}} = \sum_{B>A} \frac{Z_A Z_B}{r_{AB}}.
\end{aligned} \tag{2.14}$$

$\mathbf{T}_{\text{n}}$ is the kinetic energy of the nuclei; $\mathbf{T}_{\text{e}}$ is the the kinetic energy of the electrons. $\mathbf{V}_{\text{ee}}$ describes the electron-electron interactions with r_{ij} the distances between the electrons. $\mathbf{V}_{\text{ne}}$ are the electron-nuclei interactions, $\mathbf{V}_{\text{nn}}$ are the nuclei-nuclei interactions, r_{iA} and r_{AB} are the corresponding distance vectors. For any given geometry, the exact all-electron wave function can be expanded in a set of electronic functions, with the expansion coefficients being functions of the nuclear coordinates:

$$\Psi_{\text{tot}}\left(\mathbf{R}, \mathbf{r}\right) = \sum_{i=1}^{\infty} \Psi_{\text{n}i}(\mathbf{R})\Psi_i\left(\mathbf{R}, \mathbf{r}\right). \tag{2.15}$$

Inserting Eq. 2.15 into the Schrödinger Equation 2.13 gives

$$\sum_{i=1}^{\infty} \left(\mathbf{T}_{\text{n}} + \mathbf{H}_{\text{e}}\right) \Psi_{\text{n}i}(\mathbf{R})\Psi_i\left(\mathbf{R}, \mathbf{r}\right) = E_{\text{tot}} \sum_{i=1}^{\infty} \Psi_{\text{n}i}(\mathbf{R})\Psi_i\left(\mathbf{R}, \mathbf{r}\right). \tag{2.16}$$

Using the position space expression $\mathbf{T}_{\text{n}} = \nabla_{\text{n}}^2$, where ∇_{n}^2 implicitly includes the mass dependence, sign and summation, multiplying $\langle\Psi_j|$ from left side and integrating over the

electron coordinates leads to

$$\nabla_{\mathrm{n}}^2\Psi_{\mathrm{n}j} + E_j\Psi_{\mathrm{n}j} + \sum_i^{\infty}\left(2\left\langle\Psi_j|\nabla_{\mathrm{n}}|\Psi_i\right\rangle\left(\nabla_{\mathrm{n}}\Psi_{\mathrm{n}i}\right) + \left\langle\Psi_j|\nabla_{\mathrm{n}}^2|\Psi_i\right\rangle\Psi_{\mathrm{n}i}\right) = E_{\mathrm{tot}}\Psi_{\mathrm{n}j}. \tag{2.17}$$

The electronic wave function has been removed from the first two terms on the left hand side while the last two terms couple different electronic states. These are the first- and second-order non-adiabatic coupling elements. In the adiabatic approximation the form of the total wave function is restricted to a single electronic surface, thus all coupling terms are neglected (only terms with $i = j$ survive). Except for spatially degenerate wave functions, the diagonal first-order non-adiabatic coupling elements are zero. We obtain

$$\left(\nabla_{\mathrm{n}}^2 + E_j + \left\langle\Psi_j|\nabla_{\mathrm{n}}^2|\Psi_j\right\rangle\right)\Psi_{\mathrm{n}j} = E_{\mathrm{tot}}\Psi_{\mathrm{n}j}. \tag{2.18}$$

The third term on the left is known as the diagonal correction and is smaller than E_j by a factor roughly equal to the ratio of the electronic and nuclear mass. In the Born-Oppenheimer approximation this term is neglected, resulting in a form of the Schrödinger equation, where the electronic energy enters the nuclear part of the Schrödinger equation as a potential energy surface $\mathrm{E}_j(\mathbf{R})$.

$$\left(\mathbf{T}_{\mathrm{n}} + E_j\left(\mathbf{R}\right)\right)\Psi_{\mathrm{n}j}\left(\mathbf{R}\right) = E_{\mathrm{tot}}\Psi_{\mathrm{n}j}\left(\mathbf{R}\right) \tag{2.19}$$

Since electrons are fermions, the total electronic wave function must be anti-symmetric with respect to interchange of any two electron coordinates. This anti-symmetry can be achieved by building the wave function from Slater determinants. The first index (rows) refers to electron coordinates, the second (columns) to single electron wave functions.

$$\Phi_{\mathrm{SD}} = \frac{1}{\sqrt{N!}}\begin{vmatrix}\Phi_1(1) & \Phi_2(1) & \cdots & \Phi_N(1)\\ \Phi_1(2) & \Phi_2(2) & \cdots & \Phi_N(2)\\ \vdots & \vdots & \ddots & \vdots\\ \Phi_1(N) & \Phi_2(N) & \cdots & \Phi_N(N)\end{vmatrix}; \quad \langle\Phi_i|\Phi_j\rangle = \delta_{ij} \tag{2.20}$$

In order to derive the Hartree-Fock (HF) equations, a last approximation is made, namely that only a single determinant is considered. Assuming we know the orbitals Φ_i, we can apply such a wave function to the electronic Hamiltonian and obtain the expectation value

of the energy,

$$\begin{aligned}
\mathrm{E} &= \sum_{i}^{\mathrm{N_{elec}}} \langle \Phi_i | \mathbf{h}_i | \Phi_i \rangle + \frac{1}{2} \sum_{ij}^{\mathrm{N_{elec}}} (\langle \Phi_j | \mathbf{J}_i | \Phi_j \rangle - \langle \Phi_j | \mathbf{K}_i | \Phi_j \rangle + V_{\mathrm{nn}} \quad \text{with,} \\
\mathbf{J}_i \left| \Phi_j(2) \right\rangle &= \langle \Phi_i(1) | \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} | \Phi_i(1) \rangle \left| \Phi_j(2) \right\rangle , \\
\mathbf{K}_i \left| \Phi_j(2) \right\rangle &= \langle \Phi_i(1) | \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} | \Phi_j(1) \rangle \left| \Phi_i(2) \right\rangle .
\end{aligned} \tag{2.21}$$

In this expression $\mathbf{h}_i$ describes the motion of one electron in the field of all the nuclei and $\mathrm{V_{nn}}$ corresponds to the nuclei-nuclei coulomb repulsion term, which is constant for a given geometry. $\mathbf{J}$ is the Coulomb operator, which is a two-electron operator describing the repulsion that an electron feels caused by the other electrons in the system. $\mathbf{K}$, the exchange operator, has no classical analogon and is a consequence of the antisymmetrisation constraint for the total electron wave function. Note that this exchange of electron indices makes this operator "non-local", since the orbital j needs to be known everywhere in space for the evaluation of the operator.

We are now interested in finding a set of spin orbitals, so that the single determinant formed from these is the best possible approximation to the ground state. Following the variational principle, the best orbitals are those, which minimise the electronic energy or at least make it stationary with respect to a change in the orbitals. The variation must be carried out under the constraint that the molecular orbitals (MOs) remain orthogonal and normalised. This can be achieved by the introduction of Lagrange multipliers with the condition that a small change in the orbitals should not change the Lagrange function defined as

$$\begin{aligned}
\mathcal{L} &= E - \sum_{ij}^{\mathrm{N_{elec}}} \left(\langle \Phi_i | \Phi_j \rangle - \delta_{ij} \right) , \\
\delta\mathcal{L} &= \delta E - \sum_{ij}^{\mathrm{N_{elec}}} \left(\langle \delta\Phi_i | \Phi_j \rangle - \langle \Phi_i | \delta\Phi_j \rangle \right) .
\end{aligned} \tag{2.22}$$

The variation of the energy is expressed as

$$\delta E = \sum_{i}^{\mathrm{N_{elec}}} (\langle \delta\Phi_i | \mathbf{h}_i | \Phi_i \rangle + \langle \Phi_i | \mathbf{h}_i | \delta\Phi_i \rangle) + \sum_{ij}^{\mathrm{N_{elec}}} (\langle \delta\Phi_i | \mathbf{J}_j - \mathbf{K}_j | \Phi_i \rangle + \langle \Phi_i | \mathbf{J}_j - \mathbf{K}_j | \delta\Phi_i \rangle ,$$

by inserting the Fock operator $\mathbf{F}_i = \mathbf{h}_i + \sum_{j}^{\mathrm{N_{elec}}} (\mathbf{J}_j - \mathbf{K}_j)$ we obtain, (2.23)

$$\delta E = \sum_{i}^{\mathrm{N_{elec}}} (\langle \delta\Phi_i | \mathbf{F}_i | \Phi_i \rangle + \langle \Phi_i | \mathbf{F}_i | \delta\Phi_i \rangle).$$

Inserting Eq. 2.22 into Eq. 2.23 and using conjugate complex properties, we obtain

$$\delta\mathcal{L} = \sum_{i}^{\mathrm{N_{elec}}} \langle \delta\Phi_i | \mathbf{F}_i | \Phi_i \rangle - \sum_{ij}^{\mathrm{N_{elec}}} \lambda_{ij} \langle \delta\Phi_i | \Phi_j \rangle + \sum_{i}^{\mathrm{N_{elec}}} \langle \delta\Phi_i | \mathbf{F}_i | \Phi_i \rangle^* - \sum_{ij}^{\mathrm{N_{elec}}} \lambda_{ij} \langle \delta\Phi_j | \Phi_i \rangle^* = 0. \quad (2.24)$$

The variation of either $\langle \delta\Phi |$ or $\langle \delta\Phi |^*$ should make $\delta\mathcal{L} = 0$, leading to the condition $\lambda_{ij} = \lambda_{ji}^*$, which means that the Lagrange multipliers are elements of a Hermitian matrix. The final set of HF equations can be written as

$$\mathbf{F}_i \Phi_i = \sum_{j}^{\mathrm{N_{elec}}} \lambda_{ij} \Phi_j, \quad (2.25)$$

or, after a unitary transformation diagonalising the matrix of Lagrange multipliers ($\lambda_{ij} = 0$ and $\lambda_{ii} = \epsilon_i$),

$$\mathbf{F}_i \Phi_i^{'} = \epsilon_i \Phi_i^{'}. \quad (2.26)$$

This special set of MOs ($\Phi^{'}$) obtained after this final rotation is called canonical MOs. The Lagrange multipliers, now reduced to the diagonal elements ϵ_i, have the physical meaning of MO energies, i.e. they are expectation values of the Fock operator in the MO basis.

2.2.2 Density functional theory

The basis for density functional theory (DFT) is the proof by Hohenberg and Kohn [36] that the ground state electronic energy is determined completely by the electron density. Hence, the goal of DFT methods is to find a suitable functional, which maps an electron density uniquely onto an energy. Early attempts of DFT models tried to express all energy components as a functional of the electron density, but these methods showed a poor

performance, lacking the ability to describe molecular binding. This is due to the fact that not only the correlation and exchange energy need to be approximated but also the kinetic part. Therefore, the wave function formalism has been partly re-introduced, leading to the Kohn-Sham (KS) theory [37].
The main idea of the KS formalism is to split the kinetic energy functional into two parts, one, which can be calculated exactly, and a small remaining correction term. This beneficial approach to the problematic kinetic energy comes at the price of reintroducing orbitals, which increases the complexity from 3 to $3N$ variables. The KS model is closely related to the HF method, sharing identical formulas for the kinetic energy, electron-nuclei and electron-electron Coulomb energies. Following the nomenclature of F. Jensen [35], the total energy in Kohn-Sham theory is given as

$$\mathrm{E}[\rho] = \mathrm{T_s}[\rho] + \int dr\ v_{\mathrm{ext}}(\mathbf{r})\rho(\mathbf{r}) + \mathrm{V_{ee}}[\rho] + \mathrm{E_{xc}}[\rho]. \tag{2.27}$$

$\mathrm{T_S}$ is the KS kinetic energy, which can be expressed in terms of Kohn-Sham orbitals as

$$\begin{aligned} \mathrm{T_s}[\rho] &= \sum_i^{\mathrm{N_{elec}}} \langle\Phi_i| - \frac{1}{2}\nabla^2|\Phi_i\rangle\,, \quad \text{with} \\ \rho(\mathbf{r}) &= \sum_i^{\mathrm{N_{elec}}} |\Phi_i|^2. \end{aligned} \tag{2.28}$$

The external electrostatic potential caused by the nuclei is denoted as v_{ext}; $\mathrm{V_{ee}}$ is the electron-electron Coulomb energy, and $\mathrm{E_{xc}}$ is the exchange-correlation energy. The exchange-correlation also accounts for a part of the kinetic energy, which is not covered by Eq. 2.28 due to the assumption of non-interacting orbitals. The KS potential is found by varying this energy expression with respect to the orbitals.

$$v_{\mathrm{eff}}(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta \mathrm{E_{xc}}[\rho]}{\delta\rho(\mathbf{r})}, \tag{2.29}$$

The KS equations are derived by inserting Eq. 2.29 into the Schrödinger Equation 2.13, which leads to

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\mathrm{eff}}(\mathbf{r})\right)\Phi_i(\mathbf{r}) = \epsilon_i\Phi_i(\mathbf{r}). \tag{2.30}$$

Density functional theory approaches

The exchange and correlation energies are the only unknown parts of the energy expression in Eq. 2.27, and the task is to find proper approximations to the correct exchange-correlation functional. The exchange energy can be considered as a quantum correction to the Coulomb repulsion, leading to an interaction between electrons with the same spin, which is already present in HF methods, but must be incorporated into DFT. In addition, there is a dynamical effect where electrons tend to avoid each other more than given by an HF wave function. This "correlation energy", which is by definition missing in the HF theory, can also be included in the exchange-correlation functional, making DFT a post-HF technique as well. Following Perdew [38], who assigned density functionals to various rungs of a ladder, the first four steps of the Jacob's ladder of DFT are:

1. local density approximation (LDA),
2. generalised gradient approximation (GGA),
3. meta-GGA,
4. hyper-GGA.

Local density approximation

In the LDA it is assumed that the exchange-correlation functional is only depending on the values of the electron density ρ at each point in space, or equivalently, that the density is only a slowly varying function. In the most successful approach based on the homogeneous electron gas, the exchange energy is given by the Dirac formula,

$$\mathrm{E_x}^{\mathrm{LDA}}[\rho] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\int \rho(\mathbf{r})^{4/3} d\mathbf{r}. \tag{2.31}$$

A more general approach is the local spin density approximation (LSDA), which introduces different densities for the spins α and β:

$$\mathrm{E_x}^{\mathrm{LSDA}}[\rho] = -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\int \left(\rho_\alpha(\mathbf{r})^{4/3} + \rho_\beta(\mathbf{r})^{4/3}\right) d\mathbf{r} = \int e_{\mathrm{x}}^{\mathrm{LSDA}}\left[\rho\left(\mathbf{r}\right)\right] d\mathbf{r}. \tag{2.32}$$

For the uniform electron gas the correlation energy has been derived analytically for the high and low density limits, but is in general determined by quantum Monte Carlo methods.

The LDA underestimates the exchange energy by $\approx$ 10%, thereby creating errors that are larger than the whole correlation energy. Electron correlation is overestimated, often by a factor close to 2, and as a consequence bond strengths are overestimated, often by $\approx$ 100 kJ/mol [35]. Overall, LDA methods provide results with an accuracy similar to that obtained by wave mechanics HF methods and have been used extensively for describing extended systems, such as metals, where the approximation of a slowly varying electron density is acceptable. Common functionals are: VWN [39], PZ81 [40], CP [41] and PW92 [42].

Generalised gradient approximation

For a further improvement a nonlocal density functional must be considered. One way to do this is to assume that the exchange and correlation energies not only dependent on the electron density, but also on its first derivatives. While including the first-order exchange term improves the exchange energy, introducing the first order correlation correction often makes the correlation energy positive and thus a straightforward inclusion leads to a model, which performs worse than the LSDA approach. The main reason for the failure is that integrating the Coulomb and Fermi holes does not yield the required values 0 and -1. In the generalised gradient approximation (GGA) this requirement is fullfilled additionally to the inclusion of the gradient of the density. The functional form of the GGA exchange and correlation energies is given by [43]:

$$\mathrm{E}_{\mathrm{x/c}}^{\mathrm{GGA}} = \int e_{\mathrm{x/c}}^{\mathrm{LSDA}} \left[\rho\left(\mathbf{r}\right)\right] g_{\mathrm{x/c}} \left[s\left(\mathbf{r}\right)\right] d\mathbf{r}; \qquad s\left(\mathbf{r}\right) = \frac{\nabla\rho\left(\mathbf{r}\right)}{\rho\left(\mathbf{r}\right)^{4/3}}. \tag{2.33}$$

$g_{\mathrm{x/c}}$ is an inhomogeneity correction factor that depends on the reduced spin density $s\left(\mathbf{r}\right)$. GGA methods are also sometimes referred to as non-local methods. Common GGA functionals are: VV10 [44], BLYP [45, 46] , PW91 [47], PBE [48] and B97-D [49].

Meta-GGA

The logical extension of GGA is to include higher order derivatives of the electron density into exchange and correlation energy functionals. The second order derivative, the Laplacian $\Delta\rho$, is connected to the orbital kinetic energy density $\tau = \sum_{i}^{\mathrm{occ.}} |\nabla\Phi_i|^2$ via the orbitals and the KS potential in Eq. 2.29. Functionals that include the kinetic energy density are

referred as meta-GGA functionals. Examples for functionals are B98 [50], OPTX [51], M06-L [52] and B97M-V [43], the functional used in this work.

Hyper-GGA

Hyper-GGA functionals, often referred as hybrid functionals, evaluate a fraction of the exchange energy using a single determinant composed of KS orbitals as a wave function in a Hartree-Fock calculation. This leads to an explicit dependence on the occupied KS orbitals and a significant improve over LDA and GGA functionals. The most famous hybrid functional is B3LYP [45, 53], which has the functional form:

$$\mathrm{E}_{\mathrm{xc}}^{\mathrm{B3LYP}} = (1-a)\mathrm{E}_{\mathrm{x}}^{\mathrm{LSDA}} + a\mathrm{E}_{\mathrm{x}}^{\mathrm{exact}} + b\Delta\mathrm{E}_{\mathrm{x}}^{\mathrm{B88}} + (1-c)\mathrm{E}_{\mathrm{c}}^{\mathrm{LSDA}} + c\mathrm{E}_{\mathrm{c}}^{\mathrm{LYP}} \tag{2.34}$$

The parameters a,b and c are fitted empirically to experimental data.

2.2.3 The B97M-V functional

The functional applied to all systems in this work is the meta-GGA functional B97M-V [43]. The exchange correlation energy has the form

$$\mathrm{E}_{\mathrm{xc}}^{\mathrm{B97M-V}} = \mathrm{E}_{\mathrm{x}}^{\mathrm{B97M-V}} + \mathrm{E}_{\mathrm{css}}^{\mathrm{B97M-V}} + \mathrm{E}_{\mathrm{cos}}^{\mathrm{B97M-V}} + \mathrm{E}_{\mathrm{NLC}}^{\mathrm{VV10}}. \tag{2.35}$$

The abbreviations used are: x for exchange, c for correlation, ss for same spin, os for opposite spin and NLC for non-local correlation. The exchange energy is calculated by multiplying the integrand of the LSDA functional with a power series correction. $\mathrm{E}_{\mathrm{cos}}$ and $\mathrm{E}_{\mathrm{css}}$ are derived from the PW92 [39, 42, 54] functional. The nonlocal correlation is taken into account via the VV10 functional [44]. A strength of B97M-V is its good prediction of non-bonded interactions and, for a local density functional, also good performance in terms of thermochemistry, although it is outperformed by the more costly hybrid functionals in this. Due to the lack of exact exchange the main weakness is predicting reaction barrier heights and systems that exhibit strong correlations. According to Ref. [55] and [43] other functionals with good non-covalent performance are ωB97X-V/D [56, 57], B97-D2 [58], LC-VV10 [44]and B3LYP-NL [59].

2.2.4 Comparison of different quantum chemical methods

Besides the already introduced quantum chemical methods there exist a variety of post-HF methods with higher accuracy but also increasing computational effort. The Rayleigh-Schrödinger perturbation theory applied to the many-electron problem in molecules leads to the Møller-Plesset (MPn) series, where n defines the order of the corrections. HF already incorporates the first order correction, thus MP2 is the most common correction applied to moderately sized systems. MP2 is known to overbind slightly, which is often counter-corrected by MP3. The oscillating behaviour of the MPn series and the high computational cost for large n has led to alternative developments. CI* are configuration interaction methods in which the wave function is constructed from a linear combination of determinants build out of Hartree-Fock orbitals. The additional characters S,D,T...indicate that singly, doubly, triply, relative to the HF configuration, excited determinants are incorporated. CC* refers to coupled cluster methods. The idea of coupled cluster methods is to include all corrections of a given type e.g. doubles to infinite order. CCSD is the "gold standard" in quantum chemistry and is used for example to optimise density functionals. In Tab. 2.2 the theoretical scaling of different quantum chemical methods is listed. The characters in brackets symbolise that these excitations are calculated by perturbation theory. This is the

Table 2.2: The scaling of the computational effort with basis set size tabulated for various quantum chemistry methods

Scaling	Method
$< N^4$	DFT w/o HF exchange
N^4	HF, DFT with HF exchange
N^5	CIS, MP2
N^6	CISD, MP3, CCSD
N^7	MP4, CC3, CCSD(T)
N^8	CISDT, MP5, CCSDT
N^9	MP6
N^{10}	CISDTQ, MP7, CCSDTQ

formal scaling in the large system limit. Note that the actual scaling is typically lower than the listed value as not all parts of the computation have the same scaling. Furthermore, it does not determine the total computation time as different methods may have large prefactors, e.g. DFT calculations have formally a lower scaling than HF, but many functionals need longer computation times. In terms of accuracy with a medium-sized basis set the

following order can often be observed [35]:

HF << MP2 < CISD < MP4 ∼ CCSD < CCSD(T)

2.3 Quantum mechanics/molecular mechanics

As an alternative to pure ab initio or empirical methods combinations of both approaches, quantum mechanics/molecular mechanics (QMMM) techniques, have been developed. These methods treat some parts of the molecular system on a quantum chemistry level while other parts are handled with FF methods. The reason for this is that systems, which would be too large for pure ab initio calculations, can still be investigated with the high accuracy of quantum chemical methods at relevant locations, while the remaining part is treated at lower computational cost employing force field approaches. However, this mixture of methods also raises a series of questions:

- Which parts of the system are treated quantum chemically and which ones classically?
- How to treat a chemical bond between the ab initio and the empirical part?
- How to treat interactions between the two subsystems, such as electrostatic interactions where FFs usually use point charges while ab initio methods use the electron distribution?

The answer to the first question depends on the system and has to be tested individually. The results may depend strongly on the partitioning. Regarding the second question, where the ab initio and the FF region are connected by a chemical bond, a monovalent linking atom needs to be introduced along the bond axis to saturate this part for the ab initio calculations. A straightforward cut through a bond would create one or more unpaired electrons in the QM subsystem. This linker atom is only present for the QM calculation but invisible to the MM atoms. Another approach is to replace the chemical bond between the QM and the MM subsystem by a double-occupied molecular orbital. The third question is another tricky problem as most force fields do not include polarisability, which naturally occurs in the QM region. This can lead to an imbalance and increase the error [61].

2.4 Stereoisomerism

Isomers are molecules with the same chemical formula but different structure. There

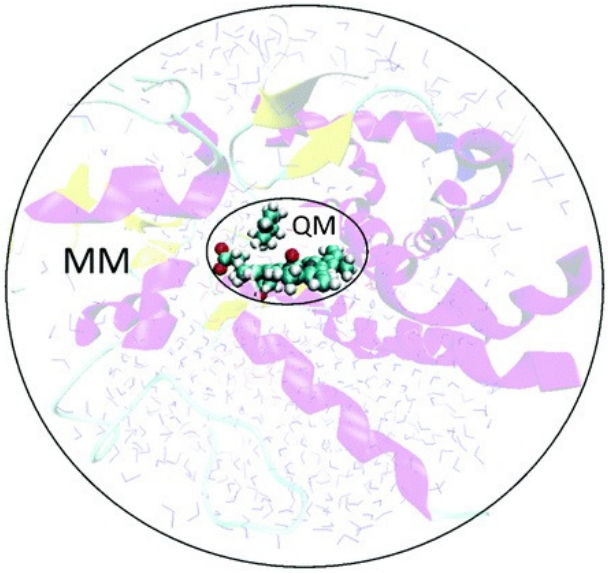


Figure 2.1: Schematic illustration of the quantum mechanical and the molecular mechanical regions of a QM/MM simulation [60].

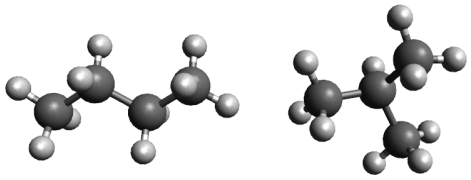


Figure 2.2: Two constitutional isomers with chemical formula C_6H_{10}

are two main types of isomerism: structural isomerism and stereoisomerism. Structural isomers (see Fig. 2.2) are molecules that are built of the same atoms and functional groups but in different order. Stereoisomers are composed out of the same atoms in the same order but in a different spatial arrangement. Examples for stereoisomers are cis-trans-isomers, conformers such as boat/chair shapes of rings and rotamers, molecules, which can be converted into each other by rotating around a single bond. Another type of stereoisomers, which is of huge importance in organic chemistry are enantiomers [62].

2.4.1 Chirality

Chirality, from the Greek word $\chi\epsilon\iota\rho$ for hand, is a geometric property of some molecules. A chiral molecule is non-superposable onto its mirror image, like left and right hand. These two forms, i.e. a chiral molecule and its mirror image, are called enantiomers. A complex formed by two enantiomers is called a diastereomer. Many things in our daily

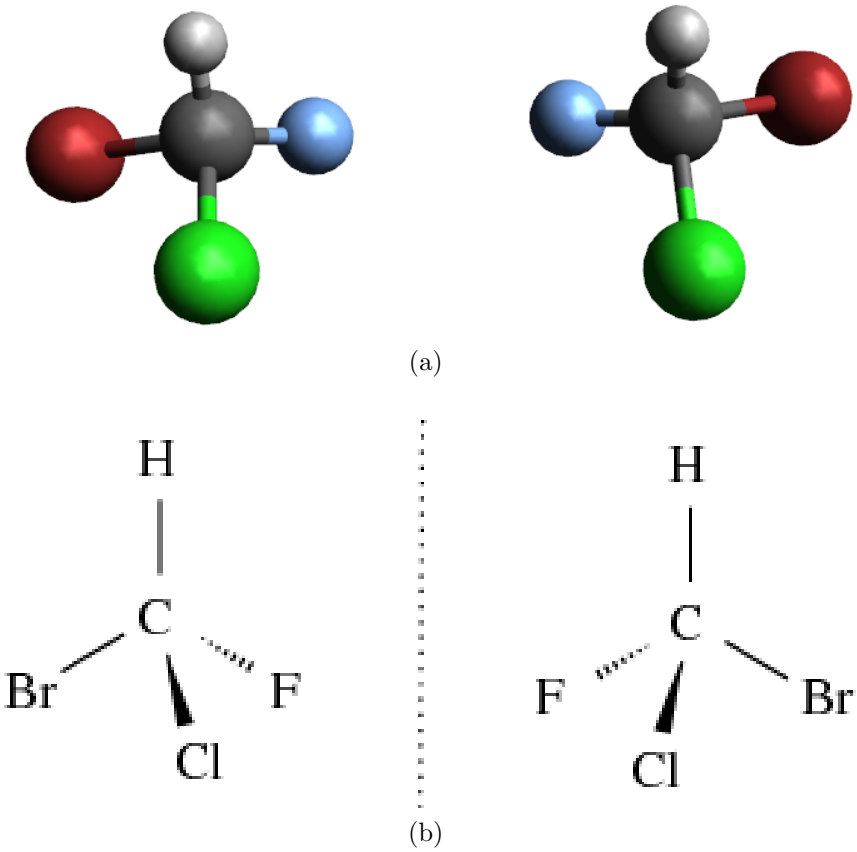


Figure 2.3: Image and mirror image are non-superposable for a chiral molecule

life are chiral such as golf rackets, shoes, corkscrews or pencil sharpeners. The chirality is based on the lack of an improper axis of rotation (S_n), which includes planes of symmetry and an inversion centre. The inversion centre is called stereogenic centre and has four different substituents. In organic chemistry the stereogenic centre is usually a carbon atom. The configurations of enantiomers are classified as dextrorotatory (R, D, or (+)) or laevorotatory (S, L or (−)). The R/S classification uses the Cahn-Ingold-Prelog priority rules. In this system, the molecule is aligned in a way in which the least important chemical group points away from the observer. Then the order of the substituents, starting from the highest priority group, is identified. If this order is clockwise, the molecule is termed the R enantiomer, if it is counter-clockwise, the molecule is denoted as S enantiomer. Another classification can be made by the only physical property that distinguishes two enantiomers, their optical activity. The (+)-form rotates the plane of a beam of linearly polarised light clockwise, the (−)-form rotates it counter-clockwise. The D-L nomenclature is derived from the Fischer projection and was largely replaced by the more accurate Cahn-Ingold-Prelog

convention. It is only used for some saccharides and amino acids [62].

Pharmaceutical importance

Interestingly, nature typically prefers one type of enantiomer over the other, although both enantiomers appear in equal amounts during the synthesis of a chemical compound. A possible explanation for this is given by the Vester-Ulbricht hypothesis [63], which states that electrons originating from beta decay are more likely to break bonds in organic molecules if spin and chirality correlate. Electrons emitted through beta decay have more often a left-handed spin, which preferable breaks the bonds in left-handed DNA. This might be the reason, why only right-handed DNA occurs in nature.
A mixture of enantiomers is called racemate. For some molecules, techniques for an asymmetric synthesis have been developed, which yield only one enantiomeric form. Although enantiomers itself have nearly the same chemical and physical properties, their biological activity with other chiral compounds may be different. The enzyme lactate dehydrogenase only oxidises (+)-lactic acid to pyruvate but not the (−)-form, because it is chiral itself and able to distinguish between right and left handed lactic acid molecules. This selectivity has a considerable impact on the biochemistry of living organisms. Many enantiomers act differently in biological systems; one enantiomer can be curing while the other is ineffective, as it is, for example, the case for adrenaline: (−)-adrenaline is heart stimulating, while (+)-adrenaline does not have any effect [64]. Obviously, separating enantiomers is therefore an important task in the pharmaceutical industry [3].

2.4.2 Methods for chiral resolution

Chiral resolution is the process of separating a racemate into its enantiomers. The basic principle of most methods is to convert the enantiomers to diastereomers, which have different achiral properties, separate them and then reconvert them to enantiomers. The first chiral resolution was performed by Louis Pasteur, who discovered that tartaric acid forms enantiopure crystals and separated them manually under the microscope. However, this method is only of historical relevance [62].

Stimulated crystallisation

The principle of this technique is to add an enantiopure compound as a nucleation site to a racemate, which then forms diastereomeric crystals. In many cases it is an acid and a base that react with each other to form a salt crystal. These can be separated by exploiting their different physical or chemical properties. Often chiral biological substances that occur naturally as single enantiomers, such as quinine or tartaric acid, are used as enantiopure compounds. The separation of (R,S)-3-Butin-2-amin by adding (+)-tartaric acid yields crystals of the (+)-enantiomer while the (-)-enantiomer is still dissolved. One advantage of this method is that the resulting crystals can often easily be segregated, e.g. by filtering, centrifugation or distillation, but only a fraction of all chiral components crystallise enantiopure and qualify for this [62].

Chromatography

The principle of chromatography is the separation of a mixture of chemical compounds with a stationary and a mobile phase. The mobile phase, either a gas or a liquid, is the transporting phase into which the initial mixture is injected. The stationary phase, most often a coarse-grained solid substance filled into a separation column, interacts with the compounds in the mobile phase that need to be separated. Each component interacts differently with the stationary phase, leading to varying retention times in the chromatograph.

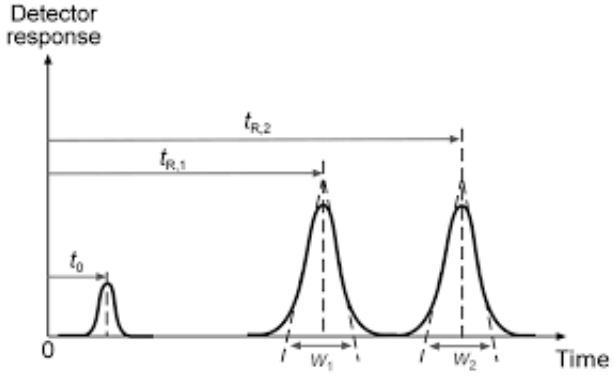


Figure 2.4: Illustration of the separation of a mixture with HPLC. After time t_0 the compound that is unaffected leaves the stationary phase. Time $t_{R,1}$ is the time needed for the first component and $t_{R,2}$ the time needed by for the second component [65].

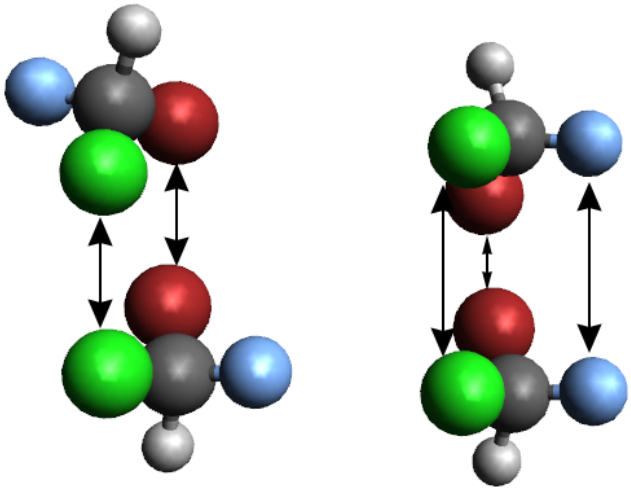


Figure 2.5: Schematic representation of the three point model. If two groups attached to a chiral centre in a molecule are exchanged, which is equal to a change from R to S chirality, only two groups can interact the same way as before.

Consequently, the compounds leave the chromatograph segregated in a timely manner as shown in Fig. 2.4. The same principle can be used to isolate a specific enantiomer from a racemic mixture. In this case, the stationary phase is coated with chiral molecules, which interact differently with components of opposite chirality, as described by the three point model, which is illustrated in Fig. 2.5. As most drug molecules are synthesised in a solution HPLC is the method of choice. HPLC uses a solid stationary phase and a liquid mobile phase. The basis of separation in chiral HPLC is the formation of temporary diastereomeric complexes within the chiral stationary phase. This causes enantiomers, which normally have identical retention times when interacting with a non-chiral stationary phase, to separate. Although the principle seems conceptually simple, the actual realisation for a given molecule is laborious and comes with several disadvantages. First, there is no universal stationary phase, which separates all types of enantiomeric pairs. Second, large amounts of solvent are required for the mobile phase in the industrial production. Another drawback is that the resolution may not suffice due to the statistical character of the separation process, leading to Gaussian-distributed signals of a certain width [3].

Porous membranes

Recently, a novel type of separation has been suggested [6], which is based on a size selection via a porous membrane. A chiral molecule, attached to the pore rim, forms a temporary complex with the free drug molecule. Theoretically also based on the 3-point model

one enantiomeric complex is slightly larger than the other. This size difference leads to an exponentially enhanced energy barrier for the transmission and thus to a separation effect. This concept was tested with a hydrogen passivated graphene pore and 1-aminoethanol as the attached, "bouncer", molecule as well as the free molecule and is illustrated in Fig. 2.6. Predicted propagation rates of $\approx 10^4$ molecules/s and a selectivity of $\approx 10^6$ for the investigated pore - free molecule combination makes this technique an interesting candidate for chiral resolution, as it would be less laborious, more ecological and better controllable than chromatography. However, a lot of problems still need to be solved, such as: up to now there are only a few synthesisable two-dimensional porous materials. How to attach the pore atom an exact position at the pore rim? Which pore-gatekeeper-free drug molecule combinations fit together and which computational techniques to predict the behaviour are favourable? A first step to answer the last question is made in this work.

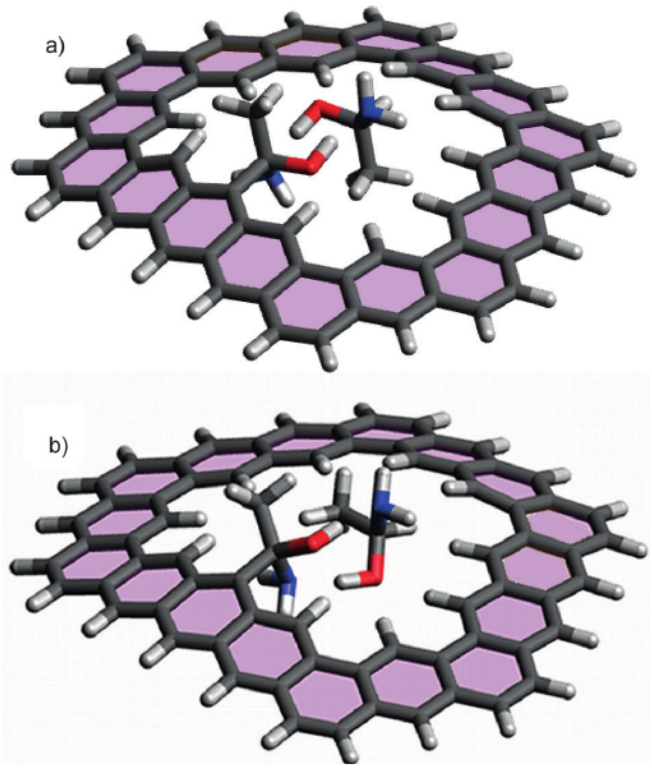


Figure 2.6: Rate-determining transition states formed by the interaction of a "bouncer"-molecule attached to the pore rim with a free molecule of a) opposite or b) same chirality. In case b) the dimer complex is slightly too large for the pore [6].

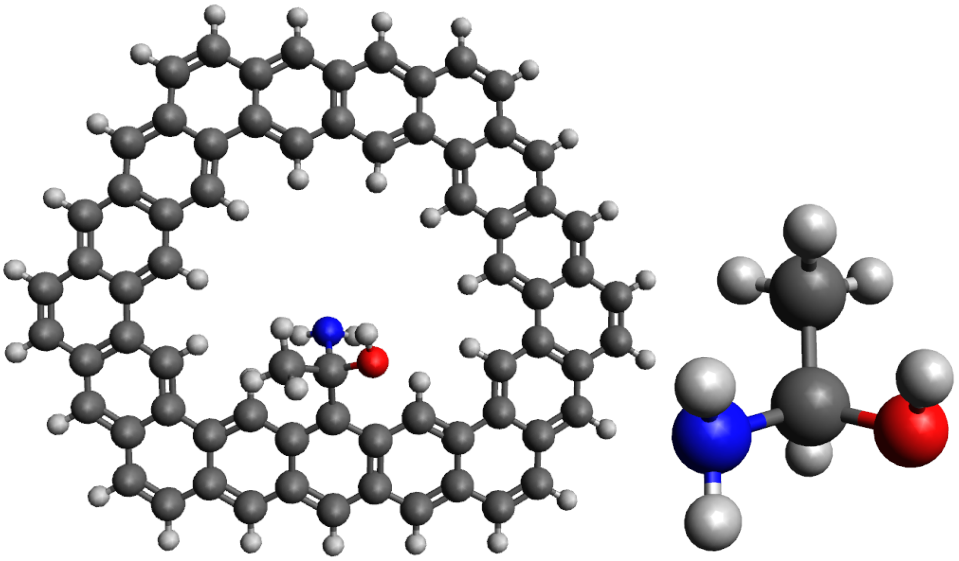
3 Membrane-based separation studies

3.1 Introduction

This chapter is dedicated to preliminary computational results of the membrane separation principle mentioned in Sec. 2.4. It aims at testing QM/MM approaches with respect to their ability to simulate the transmission of chiral molecules through a membrane. The variable determining the selectivity and transmission rate for the propagation of molecules is the energy barrier height. Therefore, the minimum energy path (MEP), including all transition states (TSs), for the transmission needs to be found. The extensive computational cost for finding such MEPs with quantum chemical methods, makes it necessary to employ other techniques, when dealing with systems of pharmaceutically relevant size. The most obvious choice is to use MM techniques. Before applying these methods, it is necessary to test them extensively in order to get knowledge about their capabilities, strengths and weaknesses. Therefore, we investigate if the chosen FFs can reproduce the ab initio data accurately, examine five chiral molecules for their possible applicability as gatekeeper - free molecule combinations and study the propagation process through a chirally unselective pore for different gas molecules.

3.2 Reproduction of quantum chemically obtained results with force field methods

A first logical step is the attempt to reproduce quantum mechanical data where available. Here we fall back on Ref. [6] and investigate the propagation of 1-aminoethanol through a graphene pore with 5 benzene rings removed and a 1-aminoethanol molecule as bouncer, as can be seen in Fig. 3.1. The MEP is determined with the nudged elastic band (NEB) method as implemented in the LAMMPS [66,67,68,69] program package for GAFF and the



(a) (b)

Figure 3.1: Test system consisting of a) functionalised pore and b) 1-aminoethanol as bouncer and drug molecule

CHARMM [70, 71, 72] suite of programs for CGenFF. The computational parameters set for all MM simulations in this work include the cutoff for the intermolecular interactions, chosen as 10 Å for GAFF, with a potential switching function from 9 Å to 10 Å and the pppm method for the long range electrostatics. For the CGenFF computations the cutoff was set to 12 Å with a force switching function from 10 Å to 12 Å for the electrostatic interactions, and a force shifting function for the van der Waals interactions.

3.2.1 CGenFF results

In Fig. 3.2 the basic structure of the transition is comparable to the ab initio results. In both cases two clear transition states and one intermediate state exist. Also, the geometry of the transition has analogies to the ab initio results. For opposite chirality of bouncer and free molecule a complex is formed, which swings through the pore, while for the same chirality the bouncer stays on one side of the pore and only the free gas molecule propagates. The transition and intermediate states for opposite chirality deviate from the ab initio results but transition state 1 for same chirality correlates with DFT results. The deformation of the pore modelled by CGenFF differs from the ab initio geometries in most cases, but the adsorption energies coincide. Unfortunately, the energy barrier height

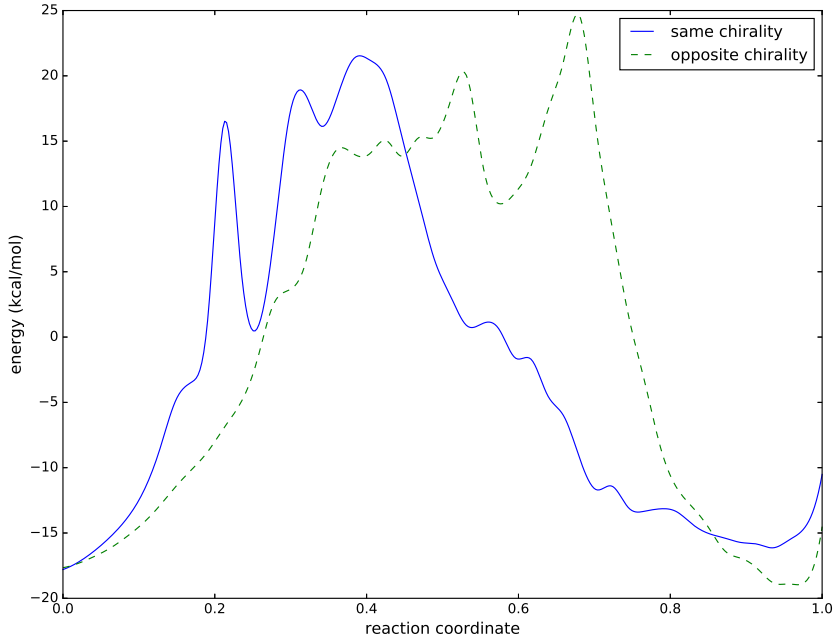
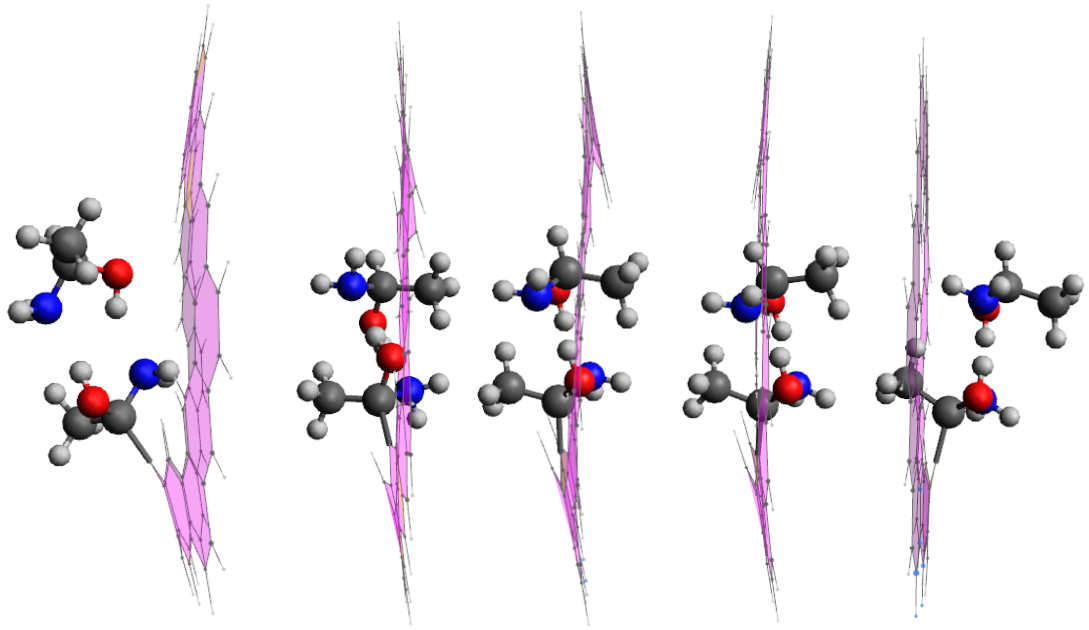


Figure 3.2: Reaction path for both enantiomers of 1-aminoethanol ascertained with NEB method and CGenFF. The continuous graph shows the propagation for the case of same chirality of bouncer and drug molecule, while the dashed graph shows the same for opposite chirality.

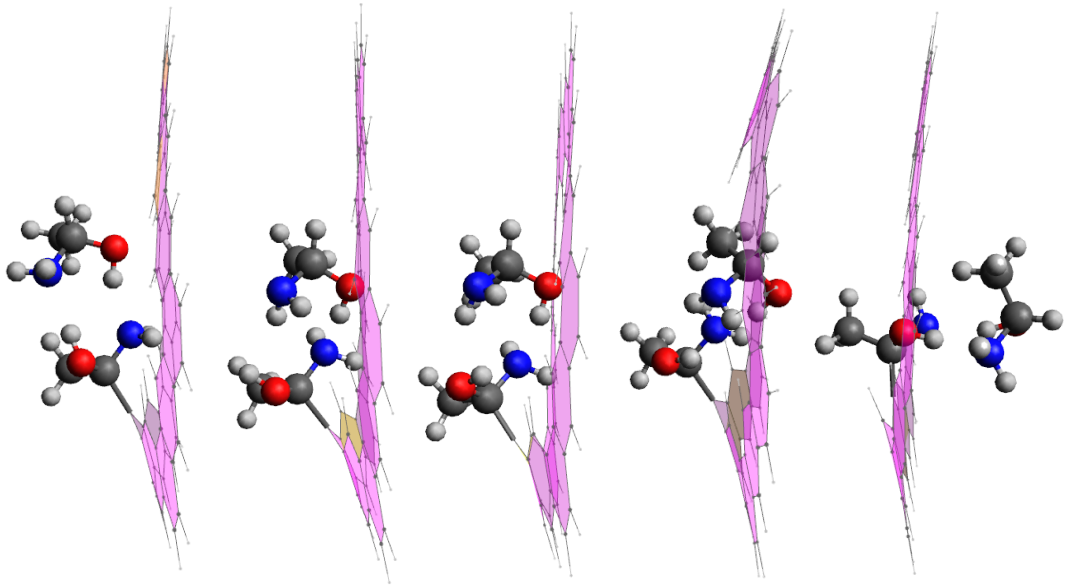
(Fig. 3.2) is about the same for both enantiomers, hence no separation effect is observable. Another backdraft is the much too high energy barrier of $\approx 22(25)$ kcal/mol, compared to $\approx -10(0)$ kcal/mol in the ab initio calculations [6]. Values in brackets are received for same chirality transition, the values in front of the brackets represent opposite chirality. This deviation is not tolerable and the CGenFF seems to be inappropriate for this approach.

3.2.2 GAFF results

However, the GAFF force field gives an even higher reaction barrier (Fig. 3.5) for one enantiomer than CGenFF, but shows a high chiral distinction in the reaction paths. Barrier heights are $8(45)$ kcal/mol compared to $-10(0)$ kcal/mol. The adsorption energies correlate with ab initio values. The propagation paths show two clear transition states and one intermediate. For opposite chirality the transition and intermediate states differ in geometry

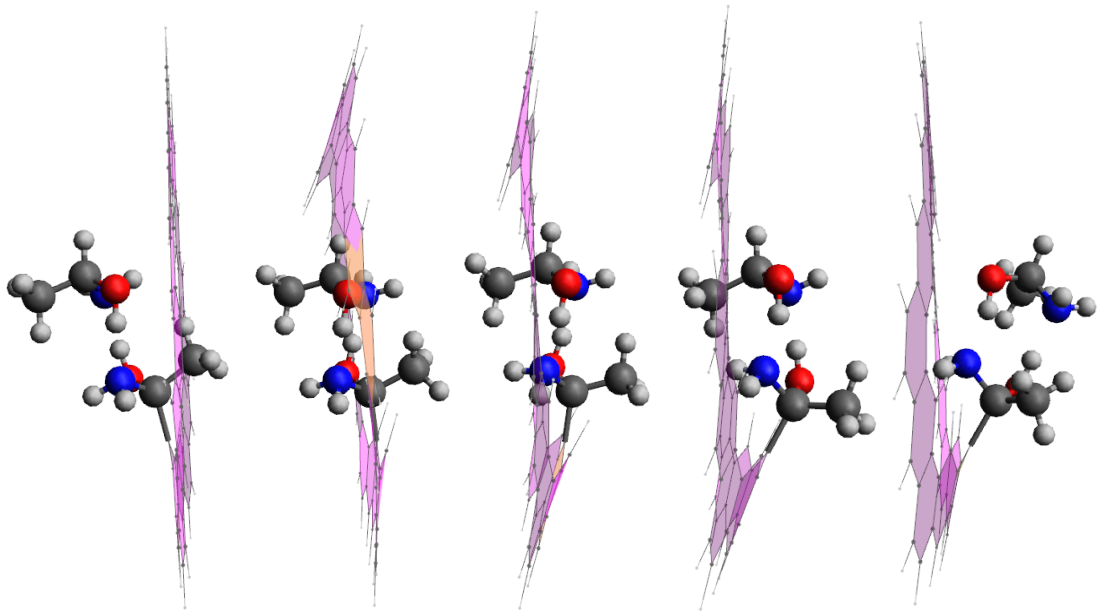


(a) left minimum (b) transition state 1 (c) intermediate (d) transition state 2 (e) right minimum

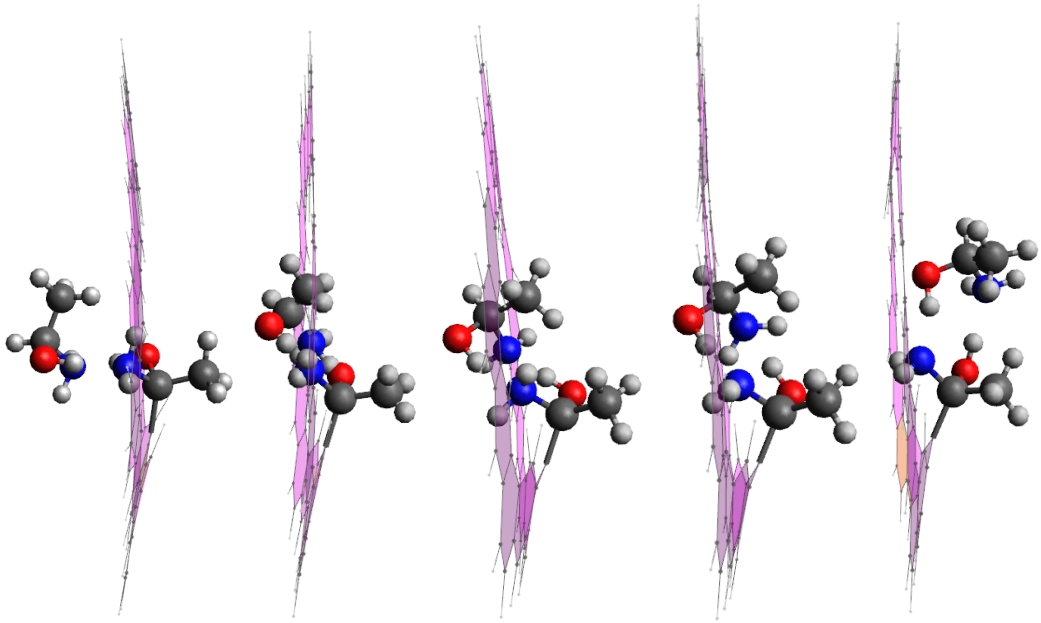


(f) left minimum (g) transition state 1 (h) intermediate (i) transition state 2 (j) right minimum

Figure 3.3: Geometries occurring along the reaction pathway for the propagation of a free gas molecule. Upper row: opposite chirality for bouncer and free gas molecule. Lower row: same chirality for bouncer and free gas molecule.



(a) left minimum (b) transition state 1 (c) intermediate (d) transition state 2 (e) right minimum



(f) left minimum (g) transition state 1 (h) intermediate (i) transition state 2 (j) right minimum

Figure 3.4: Geometries occurring along the reaction pathway for the propagation of a free gas molecule. Upper row: opposite chirality for bouncer and free gas molecule. Lower row: same chirality for bouncer and free gas molecule.

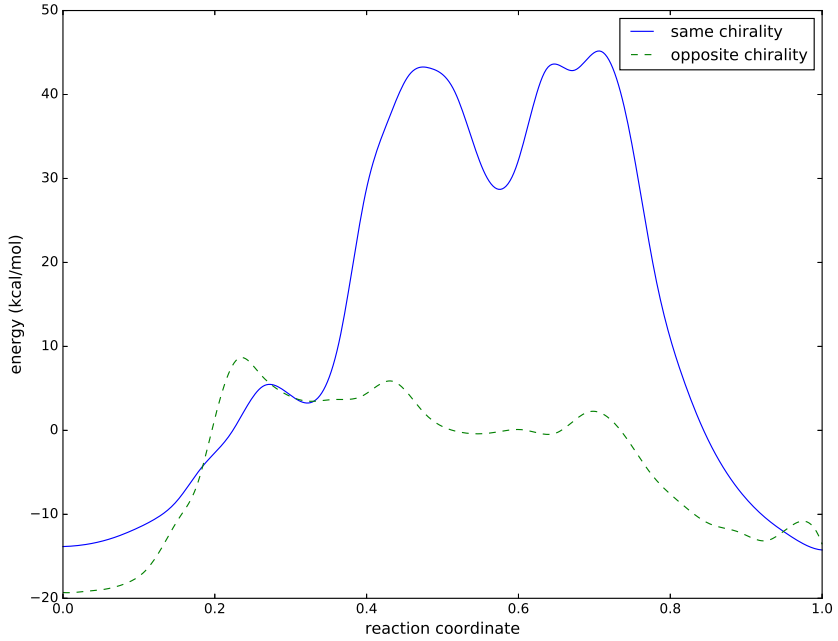
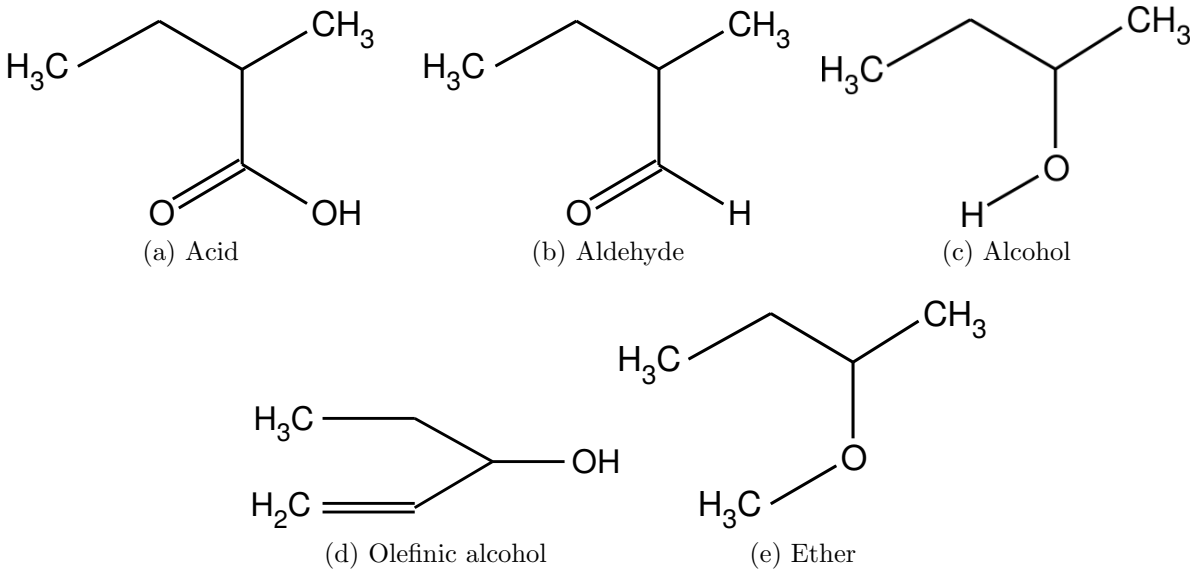


Figure 3.5: Reaction path for both enantiomers of 1-aminoethanol ascertained with NEB method and GAFF. The continuous graph shows the propagation for the case of same chirality of bouncer and drug molecule, while the dashed graph shows the same for opposite chirality.

as well as in energy from the DFT results as can be seen in Fig. 3.4. In particular, the deformation of the pore shows strong deviations. In common with the ab initio geometries is the movement of the diastereomeric complex in which the bouncer and the drug molecule swing through, while the bouncer stays on the same side for the same chirality simulation. For this case, transition state 2 and its corresponding intermediate partially correlate with the ab initio results until differences in the pore deformation occur.

3.2.3 Summary

The majority of geometries at critical points of the transmission differs from the ab initio results. Only some aspects, such as the movement of the bouncer either staying on one side or swinging through, are conserved. The shape of the reaction paths are qualitatively similar for the FFs. All propagation paths have at least two transition states and one



(a) Acid (b) Aldehyde (c) Alcohol

(d) Olefinic alcohol (e) Ether

Figure 3.6: Chiral molecules investigated

intermediate. Some minor peaks appear, which have no corresponding ab initio TSs, but have a negligible influence on the propagation process. The adsorption energies correlate with the DFT values. For CGenFF, the energy barrier height is much larger than in the ab initio computations and no chiral distinction is observable. Energy barriers obtained by GAFF are different for both enantiomers, but do still deviate vastly from the DFT values. A pure FF treatment with current force fields seems unable to capture essential features of a pore transmission. Therefore, QM/MM calculations seem inevitable.

3.3 Complex formation

As the FFs yield different results for the transitions, we split the problem into two parts, which will be investigated separately: the formation of the bouncer - drug molecule complex and the transmission through the pristine pore. This separation is necessary in order to identify if one of these parts can be modelled with one of our chosen FFs. The spatial extent of the diastereomeric complex is the crucial factor for the energy barrier of the propagation process. For a high selectivity of enantiomers the size difference of the complexes is the most important factor, more than their relative formation energies. The next step is to test if the FFs reproduce ab initio computed complex sizes within an acceptable error. In order to verify the DFT results three different functionals are employed.

The ab initio calculations are performed with the Q-Chem program package [73], using the B97M-V, the B97-D [74] and the B3LYP [45, 53] functional and the aug-cc-pVDZ [75] basis set.

Both enantiomers of five small organic molecules, see Fig. 3.6, are chosen and combined with each other. The energies of the 25 resulting complexes are minimised with DFT methods from various initial geometries to find the lowest minima. The energies and complex diameters of the lowest minima for both enantiomers are compared in Tab. 3.1 and 3.2, following Eq. 3.1 and 3.2 in order to find combinations of functional groups that are chiral distinguishable.

$$\Delta E = E_{RS} - E_{SS} \tag{3.1}$$

$$\Delta R = R_{RS} - R_{SS} \tag{3.2}$$

The subscript RS is chosen for a complex formed by molecules of different chirality, whilst the subscript SS indicates that it is formed by enantiomers of same chirality. An error of 1 kcal/mol is assumed for the energy of a DFT computation, yielding an error of 2 kcal/mol for energy differences due to the law of error propagation. An important note is that only the lowest minima are compared in the tables. However, some complexes show several optional geometries, with energies within chemical accuracy of the lowest minimum. Therefore, the tabulated size differences are only indicators for chiral distinction of the bouncer molecule, but do not necessarily result in a high chiral resolution when propagating through a functionalised pore.

3.3.1 Density functional theory results for the complexes

The comparison of the results (Tab. 3.1, Tab. 3.2) obtained by the three mentioned density functionals reveals a clear correlation between the two van der Waals corrected functionals, in energy differences, as well as in size differences. However, B3LYP results deviate strongly from the other two functionals due to the lack of a long-range correction. For most complexes the energy differences are very small and below the accuracy of the calculations except for the group of alcohol-acid complexes. The analysis of the results for the complex diameters shows that especially the aldehyde containing complexes, but also the alcohol-acid complexes, are of a different spatial extent for both combinations of enantiomers. These diameter differences are strong indicators that these combinations can be

Molecule	B97M-V	B3LYP	B97-D
Alcohol	-0.21	0.06	0.14
Acid	0.36	0.32	0.28
Aldehyde	-0.72	-0.17	0.03
Ether	1.67	2.02	1.0
Olefinic alcohol	0.78	0.93	-0.75
Alcohol-acid	-5.7	-6.58	-4.33
Alcohol-aldehyde	-0.12	-0.51	-0.96
Alcohol-ether	0.31	0.42	-0.16
Alcohol-olefinic alcohol	-0.42	0.25	-0.92
Acid-aldehyde	-0.03	0.01	-0.17
Acid-ether	1.08	1.62	0.44
Acid-olefinic alcohol	-1.8	1.41	-1.92
Aldehyde-ether	0.68	0.14	0.27
Aldehyde-olefinic alcohol	0.77	0.64	0.6
Ether-olefinic alcohol	-0.22	-0.27	-0.22

Table 3.1: Results of different density functionals for complex energy differences, according to Eq. 3.1, of the investigated chiral molecules ΔE in kcal/mol.

used as gatekeeper - free drug molecule combinations in pore propagation studies.

3.3.2 CGenFF complex results

Starting from local minima found via ab initio computations, the complexes are optimised by applying the CGenFF, GAFF and MMFF94 force fields. The results for the energy and complex size differences between both enantiomers for CGenFF are presented in Tables 3.3 and 3.4. A comparison of the ab initio and CGenFF energies does not show any correlation except for the alcohol-acid complex. However, this should not be overrated as most values are within the inaccuracy of the DFT calculations. The size differences do correlate for roughly half of the complexes, while they differ for all others. Note that only the difference between complex diameters for both enantiomers is compared. DFT and FF predictions of complex diameters can also vary for certain combinations.

Molecule	B97M-V	B3LYP	B97-D
Alcohol-alcohol	0.06	-0.29	0.04
Acid-acid	-0.11	-0.1	-0.05
Aldehyde-aldehyde	0.54	0.18	0.79
Ether-ether	-0.34	0.94	-0.41
Olefinic alcohol-olefinic alcohol	-1.04	-0.64	0.51
Alcohol-acid	1.11	0.35	1.11
Alcohol-aldehyde	-0.04	-0.35	0.35
Alcohol-ether	0.42	-0.62	0.4
Alcohol-olefinic alcohol	0.2	-1.21	-0.24
Acid-aldehyde	-1.67	0.35	-1.95
Acid-ether	-0.53	-0.2	-0.39
Acid-olefinic alcohol	-0.28	0.24	0.22
Aldehyde-ether	0.71	-0.21	-0.43
Aldehyde-olefinic alcohol	-0.35	0.56	-0.1
Ether-olefinic alcohol	-0.62	-0.47	-0.54

Table 3.2: Results of different density functionals for complex size differences, according to Eq. 3.2, of the chosen chiral molecules ΔR in Å.

Molecule	Aldehyd	Olef alc	Ether	Alcohol	Acid
Aldehyd	0.12	-0.13	0.54	-1.91	0.29
Olef alc		-1.28	-0.16	-1.51	-5.97
Ether			2.22	1.10	2.26
Alcohol				1.23	-5.44
Acid					0.0

Table 3.3: CGenFF results for complex energy differences, according to Eq. 3.1, of the chosen chiral molecules ΔE in kcal/mol.

Molecule	Aldehyd	Olef alc	Ether	Alcohol	Acid
Aldehyd	0.29	-0.03	-0.44	1.57	-0.08
Olef alc		-0.37	-0.3	-2.4	0.04
Ether			-0.28	0.36	0.34
Alcohol				-0.36	1.22
Acid					-0.01

Table 3.4: CGenFF results for complex size differences, according to Eq. 3.2, of the chosen chiral molecules ΔR in Å.

Molecule	Aldehyd	Olef alc	Ether	Alcohol	Acid
Aldehyd	-0.65	0.13	0.72	0.79	0.05
Olef alc		-1.93	-0.52	0.55	-3.49
Ether			1.17	1.53	2.89
Alcohol				0.08	-6.31
Acid					1.68

Table 3.5: GAFF results for complex energy differences, according to Eq. 3.1, of the chosen chiral molecules ΔE in kcal/mol.

Molecule	Aldehyd	Olef alc	Ether	Alcohol	Acid
Aldehyd	1.71	1.04	0.06	0.14	-0.02
Olef alc		1.3	0.14	0.24	-0.85
Ether			-1.68	0.47	0.38
Alcohol				0.07	1.41
Acid					0.05

Table 3.6: GAFF results for complex size differences, according to Eq. 3.2, of the chosen chiral molecules ΔR in Å.

3.3.3 GAFF complex results

The GAFF results (Tab. 3.5, Tab. 3.6) for the energies look very similar to those of CGenFF. Most energies correlate within the computational accuracy of the DFT calculations. The minimum geometries differ for nearly all cases from the ab initio results and GAFF performs worse than the CGenFF. Only the alcohol-acid complex energies and diameters coincide with the ab initio and CGenFF values.

3.3.4 MMFF94 complex results

Results for the complex energies of MMFF94 (Tab. 3.7, Tab. 3.8) are the same as for the other FFs, as most energy values are within the accuracy of the DFT methods. For the geometries, MMFF94 has a bad performance and only the alcohol-acid complexes seem to be described properly. For the other complexes hardly any correlation to the ab initio values can be found.

Molecule	Aldehyd	Olef alc	Ether	Alcohol	Acid
Aldehyd	-0.24	0.55	-0.17	-0.47	0.13
Olef alc		0.04	0.08	-0.16	-3.79
Ether			2.06	0.29	1.86
Alcohol				-0.58	-5.06
Acid					0.62

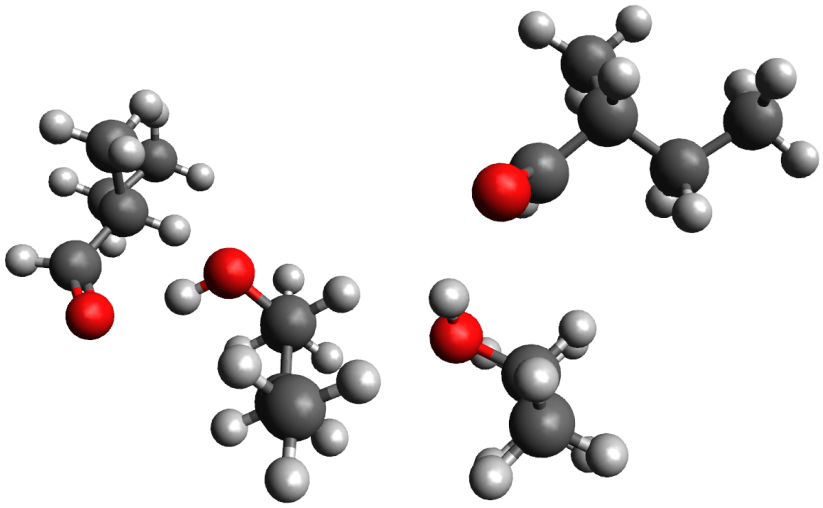
Table 3.7: MMFF94 results for complex energy differences, according to Eq. 3.1, of the chosen chiral molecules ΔE in kcal/mol.

Molecule	Aldehyd	Olef alc	Ether	Alcohol	Acid
Aldehyd	-2.38	0.21	0.26	1.49	-0.27
Olef alc		-1.48	0.14	-2.06	0.01
Ether			-0.54	0.38	-0.27
Alcohol				0.05	1.75
Acid					0.07

Table 3.8: MMFF94 results for complex size differences, according to Eq. 3.2, of the chosen chiral molecules ΔR in Å.

3.3.5 Summary of the complex formation studies

A comparison of the DFT results reveals a good agreement between different dispersion corrected functionals, which indicates reproducible and trustworthy results for the complex diameters. B3LYP, which lacks of a long-range correction produces , strongly deviating results. For most cases, the complex energy differences lie within the estimated error of the DFT computations for all force fields. However, the complex geometries show almost no correlation between the DFT and FF results for GAFF and MMFF94. CGenFF seems to perform better, but the results still differ in many cases, e.g. for the acid - olefinic alcohol complexes. This indicates that the complex formation can not be described with FFs and hence must be calculated with more accurate methods. The reason for this seems to be that the complex electron density, which determines the attraction and repulsion of the complexes, can not be modelled accurately enough with isotropic Coulomb potentials generated by point charges and Lennard-Jones potentials with fixed mixing rules. Nevertheless, some complexes, such as alcohol-acid and the aldehyde containing complexes, can be identified as potential candidates for pore propagation studies. The latter are necessary because only the knowledge of the reaction pathway can confirm the



(a) RS (b) SS

Figure 3.7: Different minimum geometries obtained for the a) RS complex b) SS complex formed by chiral aldehyde and alcohol.

chiral distinction of a pore-drug molecule system.

3.4 Pore propagation

After investigating the formation of complexes, the next step is to examine the FFs for their ability to describe the second part of the simplified system: the unfunctionalised pore. The pore usually consists of many more atoms than the complex, but these atoms do mainly contribute to the propagation via intermolecular interactions and are less included in the deformation process due to their distance from the pore rim. The aim of this study is to see if the chosen force fields can capture these chirally insensitive but yet important interactions. For a first computational calculation, we set up a series of graphene-based porous membranes.

3.4.1 System setup

The investigated systems consist of a graphene or graphene-like pore and a gas molecule. As a first candidate material we choose a graphdiyne pore, which is a triangularly patterned, two-dimensional hydrogen-carbon structure with phenyl rings at the corners and

four alternating singly-triply bonded carbon atoms as linkers in between. The second pore is a graphene pore with three benzene rings removed, which also gives the pore a triangular shape. The last examined pore is a graphene pore with four benzene rings erased, thus having a rhomboid-shaped hole. All pores are constructed totally planar, limited only to the rings needed to form them and are saturated with hydrogen atoms on the edges. This is necessary due to computational limitations. Geometries have been optimised before the propagation scans and are kept rigid during the transmission of free gas molecules. This approximation is reasonable due to the large pore diameters of 5.6 Å (graphdiyne and 3-ring pore) and 6.0 Å (4-ring pore), compared to the kinetic diameters of the studied gases: 3.46 Å (N_2), 3.64 Å (O_2) and 3.30 Å (CO_2) [76]. The molecules are positioned in steps of 0.25 Å from the geometric centre of the pore up to a distance of 4 Å, on a line perpendicular to the pore plane.

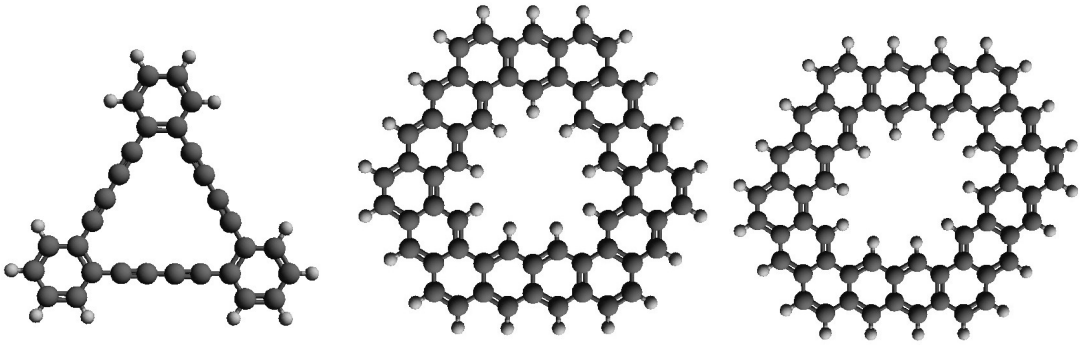
For the ab initio calculations the B97M-V, B97-D and B3LYP functionals and the cc-pVDZ basis set for the pore and the cc-pVQZ [77] basis set for the gas molecules are used. The application of two different basis sets helps to reduce the great imbalance in the number of available basis functions for the pore and the gas molecule, but a large basis set superposition errors (BSSEs) still remains. Therefore, the counterpoise correction of Boys and Bernardi [78] is applied. For the molecular mechanics computations with CGenFF, parameters for N_2 and O_2 are taken from the literature as there are none included in the force field parameter sets [79, 80]. MMFF94 lacks parameters for CO_2 and O_2 and due to the special van der Waals energy function no parameters could be found in the literature. As a consequence no transmission scans for these two gases are performed with the MMFF94 force field.

3.4.2 Comparison of propagation paths

Graphdiyne

The results for the energy path of the graphdiyne pore are interpolated with cubic splines and are plotted in Fig. 3.9, 3.10, 3.11 and 3.12.

For the nitrogen molecule (Fig. 3.9, Tab. 3.9) the results show an adsorbative behaviour for distances of around 2 Å and a repulsive potential for shorter distances. The adsorption is caused by the van der Waals attraction; the rise in the energy is explained by the Pauli repulsion due to overlapping orbitals. While the CGenFF force field and the DFT



(a) (b) (c)

Figure 3.8: Pores chosen for gas propagation studies: a) graphdiyne, b) graphene with 3 benzene rings removed and c) graphene with 4 benzene rings removed

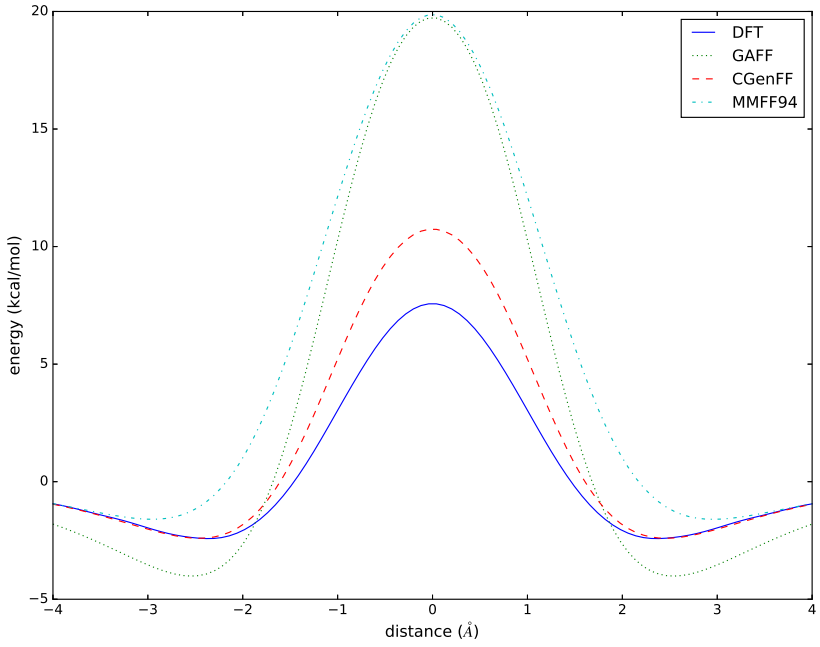


Figure 3.9: Energy path for the propagation of N_2 through the graphdiyne pore.

Table 3.9: Adsorption energy E_{ads}, energy barrier $E_{barrier}$ and the adsorption distance r_{ads} of N_2 for the propagation through the graphdiyne pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-2.4	2.25	7.6
CGenFF	-2.4	2.50	10.7
GAFF	-4.0	2.50	19.7
MMFF94	-1.6	3.00	19.9

Table 3.10: Adsorption energy, energy barrier and the adsorption distance of O_2 for the propagation through the graphdiyne pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-2.3	2.25	2.6
CGenFF	-2.8	2.25	3.9
GAFF	-4.0	2.50	4.7

calculations give similar results for greater distances, the paths begin to significantly differ for distances smaller than 2 Å thus overestimating the energy barrier by 41 percent. The MMFF94 shows a lower adsorption energy, but rises earlier and coincides with the GAFF energy at the peak, which is approximately three times the ab initio value and therefore highly unrealistic. The GAFF force field yields the highest adsorption energy.

A comparison of the results obtained with different density functionals see Fig. 3.10 reveals that there are differences concerning barrier heights and adsorption energies. While B97M-V gives the lowest ones, B97-D deviates within an acceptable error margin. B3LYP gives an energy barrier nearly twice as high and shows no attractive interaction in the investigated range.

The overall shape of the graphs for O_2 and the graphdiyne pore (Fig. 3.12, Tab. 3.10) is rather similar to the nitrogen results. The difference of the DFT and CGenFF energies is within chemical accuracy down to a distance of 1 Å. The performance of the GAFF force field is slightly better, but still largely deviating from the ab initio graph.

For CO_2 (Fig. 3.11, Tab. 3.11) the zero distance is set to the point when the carbon atom is in the middle of the pore and one oxygen is on each side. While GAFF fails terribly, CGenFF is close to the DFT results. A remarkable finding is the local minimum at zero

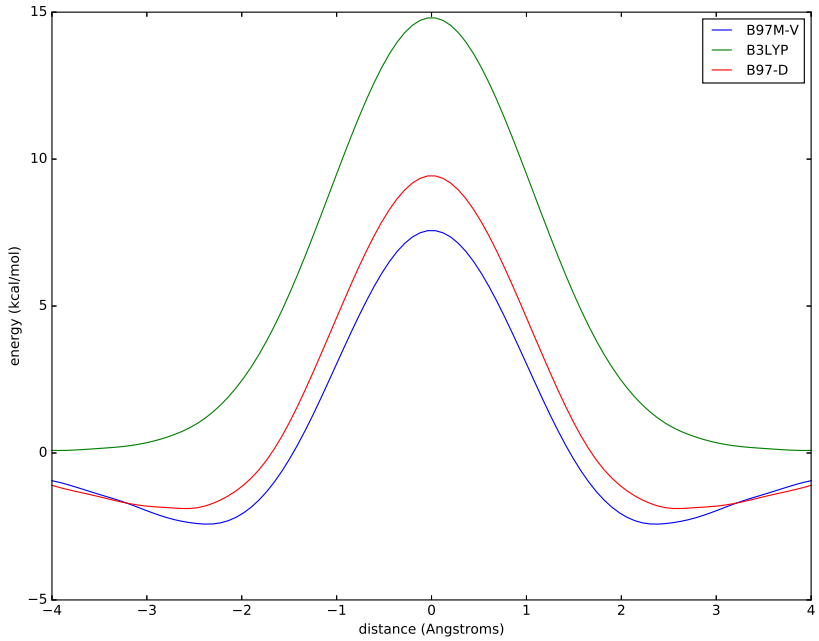


Figure 3.10: Comparison of energy paths for the propagation of N_2 through the graphdiyne pore for different density functionals.

Table 3.11: Adsorption energy, energy barrier and the adsorption distance of CO_2 for the propagation through the graphdiyne pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-2.5	2.75	1.1
CGenFF	-3.1	2.75	0.4
GAFF	-5.2	2.75	19.3

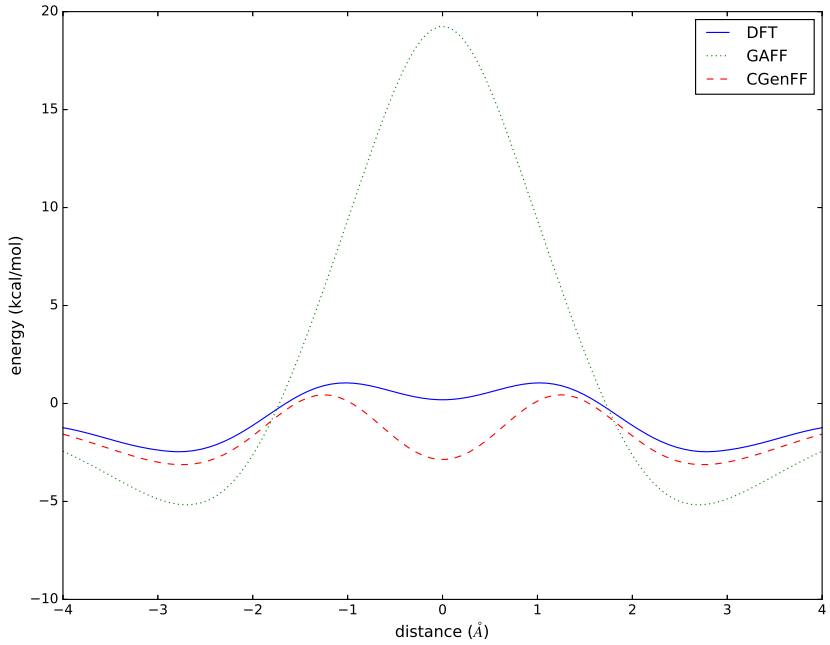


Figure 3.11: Energy path for the propagation of CO_2 through the graphdiyne pore.

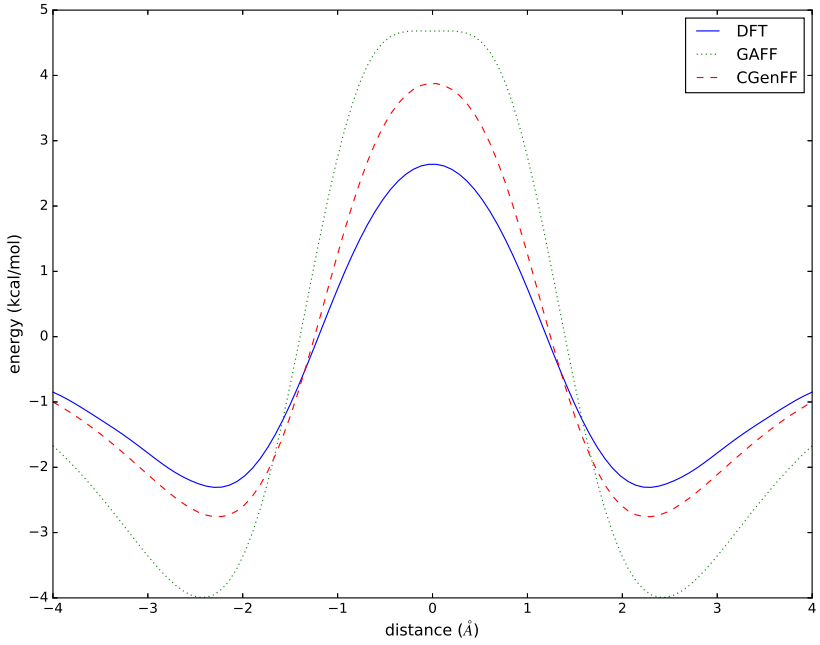


Figure 3.12: Energy path for the propagation of O_2 through the graphdiyne pore.

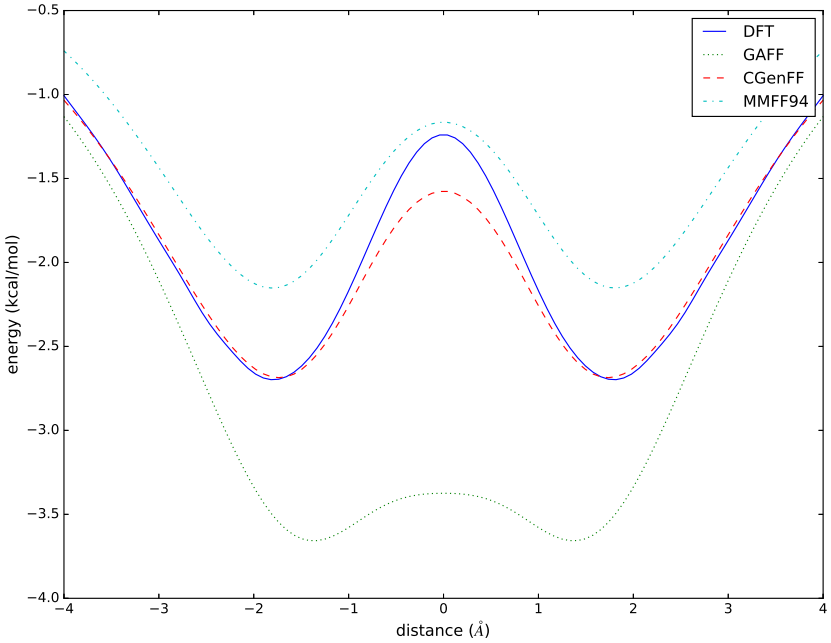


Figure 3.13: Energy path for the propagation of N_2 through the 3-ring pore.

distance, caused by the special electron density distribution of CO_2. The oxygen atoms are more electro-negative than the carbon atom, which leads to an increased electron density at both ends of the molecule. Hence, energy maxima occur whenever an oxygen atom passes the pore.

3-ring pore

For the 3-ring pore the results for the energy path are interpolated with a cubic spline and plotted in Fig. 3.13, 3.14, 3.15 and 3.16.

Due to the large diameter of the pore, which basically means that the intermolecular interactions are dominated by the attractive part of the Lennard-Jones potential, the energy for the propagation of N_2 (Fig. 3.13, Tab. 3.12) is negative throughout the whole scan. In this case, CGenFF and MMFF94 force fields give about the same results as the DFT computations. GAFF is not able to properly describe the propagation and diverges strongly at close range.

Table 3.12: Adsorption energy, energy barrier and the adsorption distance of N_2 for the propagation through the 3-ring pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-2.7	1.75	-1.2
CGenFF	-2.7	1.75	-1.6
GAFF	-3.6	1.25	-3.4
MMFF94	-2.2	1.75	-1.2

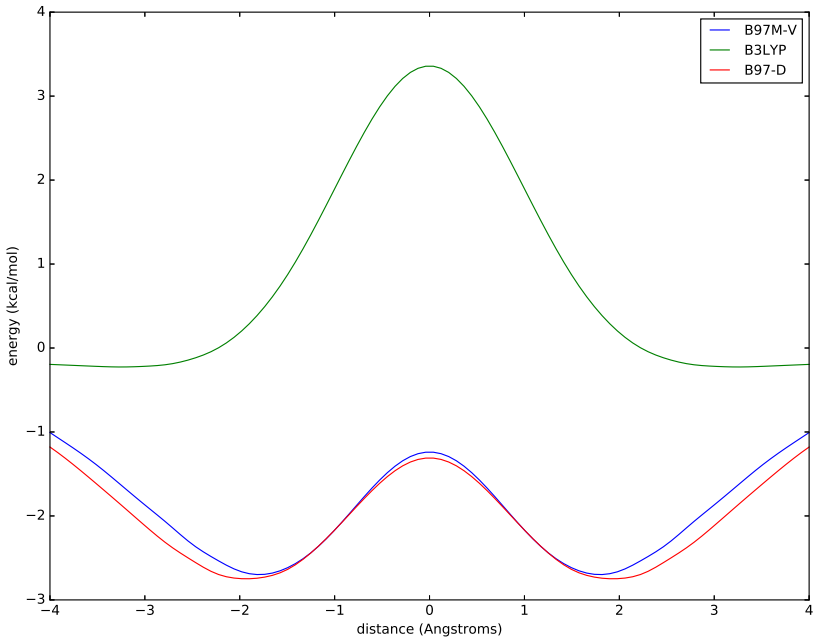


Figure 3.14: Comparison of energy paths for the propagation of N_2 through the 3-ring pore for different density functionals.

Table 3.13: Adsorption energy, energy barrier and the adsorption distance of O_2 for the propagation through the 3-ring pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-2.7	1.25	-2.2
CGenFF	-3.0	0.00	-3.0
GAFF	-4.1	0.00	-4.1

Table 3.14: Adsorption energy, energy barrier and the adsorption distance of CO_2 for the propagation through the 3-ring pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-4.3	2.0	-1.9
CGenFF	-4.1	2.25	-2.1
GAFF	-5.9	2.25	-0.6

For this pore the results of the three investigated functionals show also deviations. B97-D and B97M-V correlate and give similar results for the barrier height and adsorption minimum, while the energies obtained by B3LYP differ largely. The force fields disagree for O_2 (Fig. 3.15, Tab. 3.13) in their basic behaviour from the ab initio results. Both FFs show no local maximum in the centre of the pore as the DFT calculations do, but especially CGenFF is again very close to them. In the GAFF graph there is a small change in the curvature when the first oxygen atom passes the pore.

For carbon dioxide (Fig. 3.16, Tab. 3.14) the CGenFF and DFT graphs match within chemical accuracy. The trend of the GAFF graph coincides with the ab initio graph but the values still differ by several kcal/mol.

4-ring pore

The results for the energy path of the graphdiyne pore are interpolated with a cubic spline and plotted in Fig. 3.17, 3.18, 3.19 and 3.20. The 4-ring pore has the largest diameter of all pores, so the lowest energy barriers can be expected. For the nitrogen molecule (Fig. 3.17, Tab. 3.15) the CGenFF graph agrees well with the ab initio data. MMFF94 has no local maximum in the middle of the pore but is energetically close to the DFT results. GAFF diverges once again and has a much higher adsorption energy. Looking at the energy paths obtained by the different density functionals, reveals that B97M-V and B97-D give equal

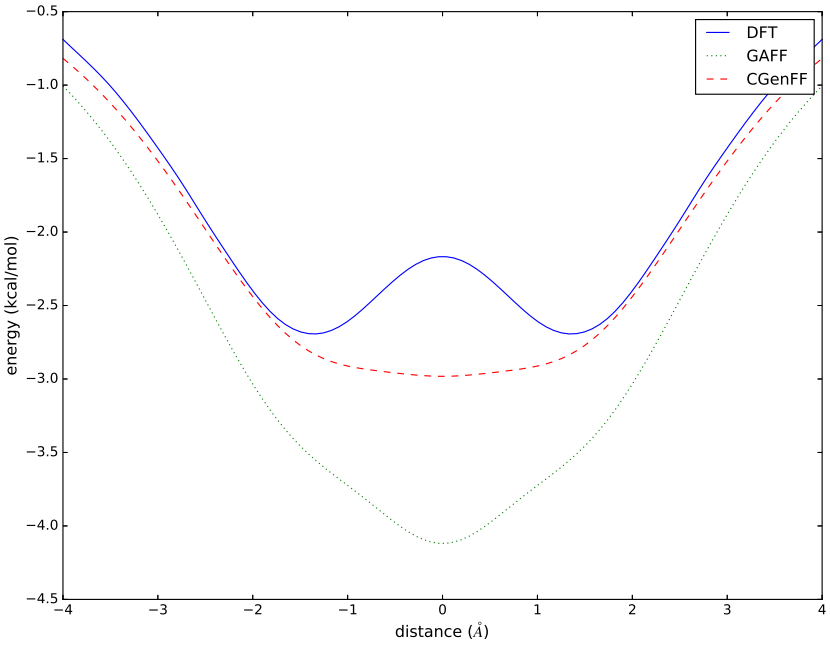


Figure 3.15: Energy path for the propagation of O_2 through the 3-ring pore.

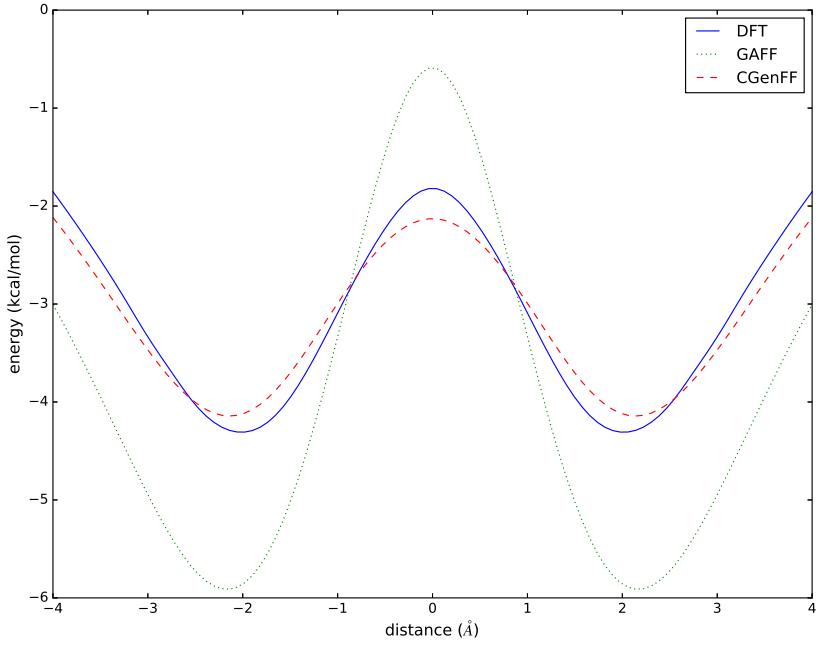


Figure 3.16: Energy path for the propagation of CO_2 through the 3-ring pore.

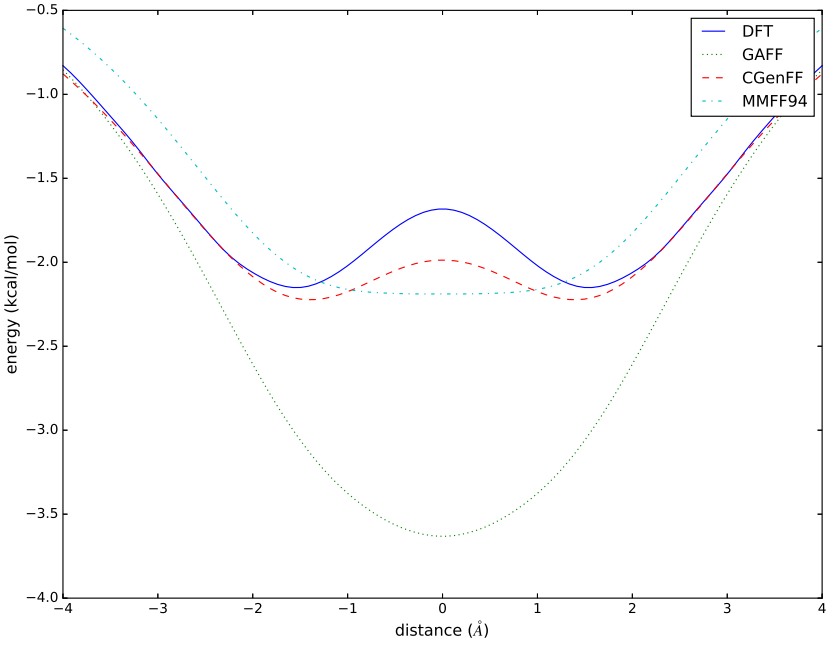


Figure 3.17: Energy path for the propagation of N_2 through the 4-ring pore.

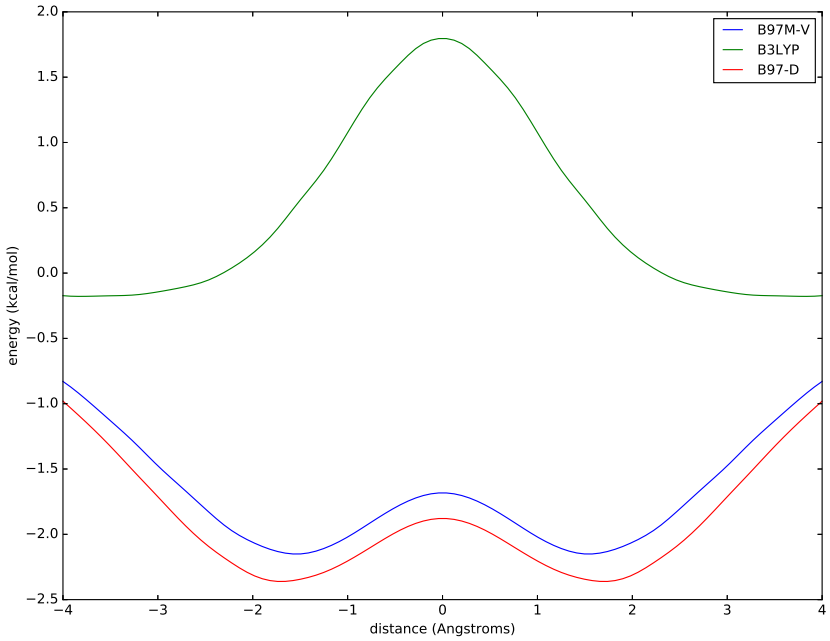


Figure 3.18: Comparison of energy paths for the propagation of N_2 through the 4-ring pore for different density functionals.

Table 3.15: Adsorption energy, energy barrier and the adsorption distance of N_2 for the propagation through the 4-ring pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-2.2	1.50	-1.7
CGenFF	-2.2	1.50	-2.0
GAFF	-3.6	0.00	-3.6
MMFF94	-2.2	0.00	-2.2

Table 3.16: Adsorption energy, energy barrier and the adsorption distance of O_2 for the propagation through the 4-ring pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-2.1	1.00	-2.0
CGenFF	-2.6	0.00	-2.6
GAFF	-3.4	0.00	-3.4

results within chemical accuracy. The third functional B3LYP differs largely concerning barrier height as well as adsorption energy and minimum energy distance.

The DFT results show a low energy barrier in the centre of the pore for molecular oxygen (Fig. 3.19, Tab. 3.16). None of the FFs can reproduce this local maximum. Again, CGenFF is closer to the ab initio results than GAFF.

The correlation between CGenFF and DFT results for CO_2 (Fig. 3.20, Tab. 3.17) is very good. Both show a distinct adsorption minimum and an energy barrier. The depth of the minimum differs moderately for GAFF while the maximum is close to the other results.

Table 3.17: Adsorption energy, energy barrier and the adsorption distance of CO_2 for the propagation through the graphdiyne pore.

Method	E_{ads} (kcal/mol)	r_{ads} (Å)	$E_{barrier}$ (kcal/mol)
DFT	-3.3	2.00	-1.5
CGenFF	-4.4	2.00	-1.9
GAFF	-3.2	2.00	-1.8

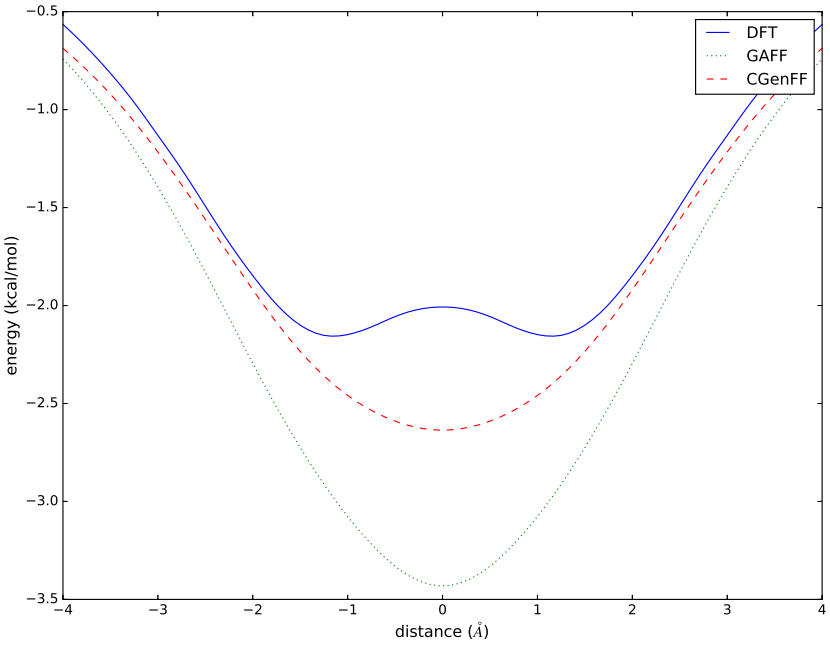


Figure 3.19: Energy path for the propagation of O_2 through the 4-ring pore.

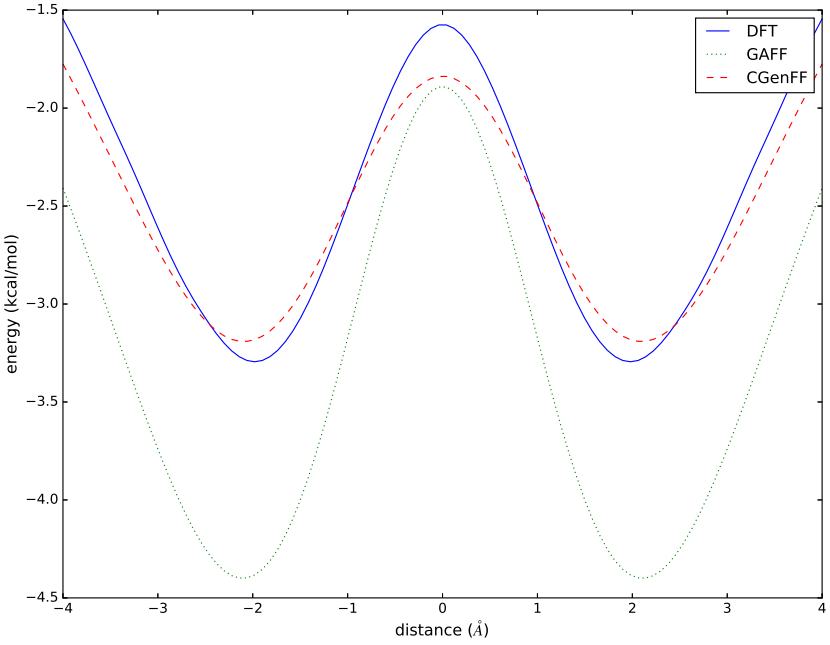


Figure 3.20: Energy path for the propagation of CO_2 through the 4-ring pore.

3.4.3 Summary

The three chosen density functionals perform as expected. B97M-V and B97-D coincide within chemical accuracy for all cases although being based on different dispersion corrections. This indicates that DFT gives reliable results if van der Waals corrected density functionals are used for propagation studies. B3LYP, which is uncorrected for these long-range interactions, fails to describe the transmission process of the gas molecules. All chosen FFs behave rather differently. Overall, CGenFF has the best performance with values lying in most cases within chemical accuracy of the DFT results. The main reasons can be identified as the availability of optimised parameters for the gases investigated and a good description of the intermolecular forces for the pore atoms. GAFF shows the worst performance. The adsorption energies are overestimated and the energy barrier heights differ significantly from the ab initio results. The reason for this is that GAFF lacks specific parameters for the investigated gases. For example, there is only one parameter set for a doubly bonded oxygen. Missing atom type definitions can therefore be identified as the main reason for inaccuracies. The MMFF94 performance is acceptable except for the graphdiyne pore. The values correlate with the ab initio results within chemical accuracy. Only for the graphdiyne pore the energy barrier is about 13 kcal/mol too high. MMFF94 has no parameters for CO_2 and O_2, which limits the significance and comparability with the other FFs. In summary, CGenFF seems to be the best FF to describe the intermolecular interactions of the pristine pore.

4 Conclusion and outlook

The reproduction of the ab initio results with molecular mechanics approaches shows large differences in energies and geometries. While the adsorption energies agree with the DFT values, the energy barrier heights differ significantly for both examined FFs. Fundamental ratios of attractive and repulsive interactions are captured by the FFs in most cases, leading to the expected pathway containing two transition states and one intermediate. Also, the geometries of important states of the reaction paths exhibit analogies to the ab initio data. Nonetheless, the different energy barrier heights and deviations of the propagation process prove that the tested force fields are not able to simulate the propagation within the accuracy needed to estimate chiral resolution performances.

The complex formation tests show hardly any correlation between ab initio and FF results, though the energy differences are within the accuracy of the DFT computations. The analysis of the different density functionals reveals that the two dispersion corrected functionals B97M-V and B97-D give coinciding results for most cases while the B3LYP functional fails to reproduce their results. The comparison of the complex geometries exposes weaknesses of all investigated FFs. CGenFF seems to be closest to ab initio data but still deviates in many cases. The other FFs perform worse and no correlation to the DFT results can be found. As high accuracy in the formation of the complex is needed, the examined FFs are inappropriate for this task. Hence, we conclude that this part of the system has to be simulated with other more accurate methods. However, some suitable combinations of gatekeeper and free drug molecules, can be found, which show significantly different complex diameters for both enantiomeric combinations e.g. the alcohol - acid dimer and the aldehyde containing complexes. One reason for this could be the high polarity and thus the differently shaped electronic density of the acid and alcohol groups combined with the repulsion of the large methyl groups. Only the investigation of the electron density of the dimers can answer this question with certainty.

The results of the gas propagation scans state that it is possible to describe the transmission

properly with CGenFF in most cases. The reason for the good performance of CGenFF is the use of optimised parameters for the gases, which are either already included in the force field or taken from the literature. GAFF and MMFF94 perform worse due to the lack of gas-specific parameters. The investigation of the performance of different density functionals shows that an explicit correction of van der Waals interactions is essential for a correct description of molecular propagation through porous membranes.

The next step will be to combine the two applied techniques for a QM/MM simulation since the complex formation needs to be simulated with quantum chemical methods while the pore can be treated with the CGenFF force field. The aim of these follow-up computations is to find the reaction paths for different drug molecule - pore combinations. Starting from these scans, combinations of functional groups for the bouncer and the free molecule can be determined, which increase the chiral selectivity at improved transmission rates. In a future step, the analysis of pharmaceutically relevant molecules and alternative porous materials such as two-dimensional metal-organic frameworks will be attempted.

5 Bibliography

[1] J. E. Rekoske, "Chiral separations," *AIChE Journal*, vol. 47, no. 1, pp. 2–5, 2001.

[2] Gübitz, Gerald and Schmid, Martin G., "Chiral separation principles in chromatographic and electromigration techniques," *Molecular Biotechnology*, vol. 32, no. 2, pp. 159–179, 2006.

[3] D. G. Watson, *Pharmaceutical Analysis: A Textbook for Pharmacy Students and Pharmaceutical Chemists, 2e.* Churchill Livingstone, 2005.

[4] T. Booth, D. Wahnon, and I. Wainer, "Is chiral recognition a three-point process?," *Chirality*, vol. 9, no. 2, pp. 96–98, 1997.

[5] A. Mesecar and D. Koshland Jr., "A new model for protein stereospecificity," *Nature*, vol. 403, no. 6770, pp. 614–615, 2000.

[6] A. W. Hauser, N. Mardirossian, J. A. Panetier, M. Head-Gordon, A. T. Bell, and P. Schwerdtfeger, "Functionalized Graphene as a Gatekeeper for Chiral Molecules: An Alternative Concept for Chiral Separation," *Angewandte Chemie International Edition*, vol. 53, pp. 9957–9960, Sept. 2014.

[7] X. Bao, L. J. Broadbelt, and R. Q. Snurr, "Computational screening of homochiral metal–organic frameworks for enantioselective adsorption," *Microporous and Mesoporous Materials*, vol. 157, no. 0, pp. 118 – 123, 2012.

[8] G. Ferey, "Hybrid porous solids: Past, present, future," *Chemical Society Reviews*, vol. 37, no. 1, pp. 191–214, 2008.

[9] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric Field Effect in Atomically Thin Carbon Films," *Science*, vol. 306, no. 5696, pp. 666–669, 2004.

[10] G. C. Paola Russo, Anming Hu, "Synthesis, properties and potential applications of porous graphene: A review," *Nano-Micro Letters*, vol. 5, pp. 260–273, 2013.

[11] L. Jiang and Z. Fan, "Design of advanced porous graphene materials: from graphene nanomesh to 3d architectures," *Nanoscale*, vol. 6, pp. 1922–1945, 2014.

[12] G. Xie, R. Yang, P. Chen, J. Zhang, X. Tian, S. Wu, J. Zhao, M. Cheng, W. Yang, D. Wang, C. He, X. Bai, D. Shi, and G. Zhang, "A general route towards defect and pore engineering in graphene," *Small*, 2014.

[13] W. Yuan, J. Chen, and G. Shi, "Nanoporous graphene materials," *Materials Today*, vol. 17, no. 2, pp. 77 – 85, 2014.

[14] V. Berry, "Impermeability of graphene and its applications," *Carbon*, vol. 62, no. 0, pp. 1 – 10, 2013.

[15] S. ei Ozeki, T. Ito, K. Uozumi, and I. Nishio, "Scanning Tunneling Microscopy of UV-Induced Gasification Reaction on Highly Oriented Pyrolytic Graphite," *Japanese Journal of Applied Physics*, vol. 35, no. Part 1, No. 6B, pp. 3772–3774, 1996.

[16] S. Huh, J. Park, Y. S. Kim, K. S. Kim, B. H. Hong, and J.-M. Nam, "UV/Ozone-Oxidized Large-Scale Graphene Platform with Large Chemical Enhancement in Surface-Enhanced Raman Scattering," *ACS Nano*, vol. 5, no. 12, pp. 9799–9806, 2011.

[17] S. P. Koenig, L. Wang, J. Pellegrino, and J. S. Bunch, "Selective molecular sieving through porous graphene," *Nature Nano*, vol. 7, pp. 728–732, 11 2012.

[18] D. C. Bell, M. C. Lemme, L. A. Stern, J. R. Williams, and C. M. Marcus, "Precision cutting and patterning of graphene with helium ions," *Nanotechnology*, vol. 20, no. 45, p. 455301, 2009.

[19] K. Celebi, J. Buchheim, R. M. Wyss, A. Droudian, P. Gasser, I. Shorubalko, J.-I. Kye, C. Lee, and H. G. Park, "Ultimate permeation across atomically thin porous graphene," *Science*, vol. 344, no. 6181, pp. 289–292, 2014.

[20] M. D. Fischbein and M. Drndić, "Electron beam nanosculpting of suspended graphene sheets," *Applied Physics Letters*, vol. 93, pp. 113107–1–3, 2008.

[21] P. Kuhn, A. Forget, D. Su, A. Thomas, and M. Antonietti, "From Microporous Regular Frameworks to Mesoporous Materials with Ultrahigh Surface Area: Dynamic Reorganization of Porous Polymer Networks," *Journal of the American Chemical Society*, vol. 130, no. 40, pp. 13333–13337, 2008.

[22] P. Xu, J. Yang, K. Wang, Z. Zhou, and P. Shen, "Porous graphene: Properties, preparation, and potential applications," *Chinese Science Bulletin*, vol. 57, no. 23,

pp. 2948–2955, 2012.

[23] K.-D. Zhang, J. Tian, D. Hanifi, Y. Zhang, A. C.-H. Sue, T.-Y. Zhou, L. Zhang, X. Zhao, Y. Liu, and Z.-T. Li, "Toward a Single-Layer Two-Dimensional Honeycomb Supramolecular Organic Framework in Water," *Journal of the American Chemical Society*, vol. 135, pp. 17913–17918, Nov. 2013.

[24] D. Frenkel and B. Smit, *Understanding Molecular Simulation, Second Edition: From Algorithms to Applications (Computational Science)*. Academic Press, 2001.

[25] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell, "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," *Journal of Computational Chemistry*, vol. 31, no. 4, pp. 671–690, 2010.

[26] W. Yu, X. He, K. Vanommeslaeghe, and A. D. MacKerell, "Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations," *Journal of Computational Chemistry*, vol. 33, p. 2451–2468, Jul 2012.

[27] K. Vanommeslaeghe and A. D. MacKerell, "Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing," *Journal of Chemical Information and Modeling*, vol. 52, p. 3144–3154, Dec 2012.

[28] K. Vanommeslaeghe, E. P. Raman, and A. D. MacKerell, "Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges," *Journal of Chemical Information and Modeling*, vol. 52, p. 3155–3168, Dec 2012.

[29] H. A. Lorentz, "Über die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase," *Annalen der Physik*, vol. 248, no. 1, p. 127–136, 1881.

[30] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *Journal of Computational Chemistry*, vol. 25, no. 9, pp. 1157–1174, 2004.

[31] J. Wang, W. Wang, P. A. Kollman, and D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations," *Journal of Molecular Graphics and Modelling*, vol. 25, p. 247–260, Oct 2006.

[32] T. A. Halgren, "Merck molecular force field. I. Basis, form, scope, parameterization,

and performance of MMFF94," *Journal of Computational Chemistry*, vol. 17, no. 5-6, pp. 490–519, 1996.

[33] E. B. W. Jr., J. Decius, and P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*. Dover Publications, 1980.

[34] T. A. Halgren, "The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters," *Journal of the American Chemical Society*, vol. 114, p. 7827–7843, Sep 1992.

[35] F. Jensen, *Introduction to Computational Chemistry*. Wiley, 2007.

[36] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Physical Reviews*, vol. 136, pp. B864–B871, Nov 1964.

[37] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review*, vol. 140, p. A1133–A1138, Nov 1965.

[38] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, "Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits," *Journal of Chemical Physics*, vol. 123, no. 6, p. 062201, 2005.

[39] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis," *Canadian Journal of Physics*, vol. 58, p. 1200–1211, Aug 1980.

[40] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," *Physical Review B*, vol. 23, p. 5048–5079, May 1981.

[41] L. A. Cole and J. P. Perdew, "Calculated electron affinities of the elements," *Physical Review A*, vol. 25, p. 1265–1271, Mar 1982.

[42] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy," *Physical Review B*, vol. 45, p. 13244–13249, Jun 1992.

[43] N. Mardirossian and M. Head-Gordon, "Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V," *Journal of Chemical Physics*, vol. 142, p. 074111, Feb 2015.

[44] O. A. Vydrov and T. Van Voorhis, "Nonlocal van der Waals density functional: The

simpler the better," *The Journal of Chemical Physics*, vol. 133, p. 244103, Dec. 2010.

[45] C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density," *Physical Review B*, vol. 37, p. 785–789, Jan 1988.

[46] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Physical Review A*, vol. 38, p. 3098–3100, Sep 1988.

[47] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation," *Physical Review B*, vol. 46, p. 6671–6687, Sep 1992.

[48] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical Review Letters*, vol. 77, p. 3865–3868, Oct 1996.

[49] S. Grimme, "Semiempirical gga-type density functional constructed with a long-range dispersion correction," *Journal of Computational Chemistry*, vol. 27, no. 15, p. 1787–1799, 2006.

[50] H. L. Schmider and A. D. Becke, "Optimized density functionals from the extended g2 test set," *The Journal of Chemical Physics*, vol. 108, p. 9624–9631, Jun 1998.

[51] W.-M. Hoe, A. J. Cohen, and N. C. Handy, "Assessment of a new local exchange functional optx," *Chemical Physics Letters*, vol. 341, p. 319–328, Jun 2001.

[52] Y. Zhao and D. G. Truhlar, "A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions," *The Journal of Chemical Physics*, vol. 125, p. 194101, Nov 2006.

[53] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *Journal of Chemical Physics*, vol. 98, no. 7, p. 5648, 1993.

[54] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," *Physical Review Letters*, vol. 45, p. 566–569, Aug 1980.

[55] N. Mardirossian and M. Head-Gordon, "Exploring the limit of accuracy for density functionals based on the generalized gradient approximation: Local, global hybrid, and range-separated hybrid functionals with and without dispersion corrections," *The Journal of Chemical Physics*, vol. 140, p. 18A527, May 2014.

[56] N. Mardirossian and M. Head-Gordon, "B97X-V: A 10-parameter, range-separated

hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy," *Physical Chemistry Chemical Physics*, vol. 16, no. 21, p. 9904, 2014.

[57] J.-D. Chai and M. Head-Gordon, "Systematic optimization of long-range corrected hybrid density functionals," *The Journal of Chemical Physics*, vol. 128, no. 8, p. 084106, 2008.

[58] P. J. Wilson, T. J. Bradley, and D. J. Tozer, "Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials," *The Journal of Chemical Physics*, vol. 115, no. 20, p. 9233, 2001.

[59] W. Hujo and S. Grimme, "Performance of the van der Waals Density Functional VV10 and (hybrid)GGA Variants for Thermochemistry and Noncovalent Interactions," *Journal of Chemical Theory and Computation*, vol. 7, p. 3866–3871, Dec 2011.

[60] R. Lonsdale, J. N. Harvey, and A. J. Mulholland, "A practical guide to modelling enzyme-catalysed reactions," *Chemical Society Reviews*, vol. 41, no. 8, p. 3025, 2012.

[61] G. Groenhof, "Introduction to QM/MM Simulations," *Biomolecular Simulations*, p. 43–66, Aug 2012.

[62] K. P. C. Vollhardt and N. E. Schore, *Organische Chemie (German Edition)*. Wiley-VCH, 2005.

[63] J. M. Dreiling and T. J. Gay, "Chirally sensitive electron-induced molecular breakup and the vester-ulbricht hypothesis," *Physical Review Letters*, vol. 113, Sep 2014.

[64] H. Hart, L. E. Craine, D. J. Hart, and C. M. Hadad, *Organische Chemie (German Edition)*. Wiley-VCH Verlag GmbH, 2007.

[65] J. Anderson, A. Berthod, V. Pino, and A. M. Stalcup, *Analytical Separation Science, 5 Volume Set*. Wiley-VCH, 2016.

[66] S. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," *Journal of Computational Physics*, vol. 117, p. 1–19, Mar 1995.

[67] G. Henkelman and H. Jónsson, "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points," *The Journal of Chemical Physics*, vol. 113, no. 22, pp. 9978–9985, 2000.

[68] G. Henkelman, B. P. Uberuaga, and H. Jonsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," *The Journal of*

Chemical Physics, vol. 113, no. 22, p. 9901, 2000.

[69] A. Nakano, "A space–time-ensemble parallel nudged elastic band algorithm for molecular kinetics simulation," *Computer Physics Communications*, vol. 178, p. 280–289, Feb 2008.

[70] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, and et al., "Charmm: The biomolecular simulation program," *Journal of Computational Chemistry*, vol. 30, p. 1545–1614, Jul 2009.

[71] H. Lee Woodcock, M. Hodošček, P. Sherwood, Y. S. Lee, H. F. Schaefer III, and B. R. Brooks, "Exploring the quantum mechanical/molecular mechanical replica path method: a pathway optimization of the chorismate to prephenate claisen rearrangement catalyzed by chorismate mutase," *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, vol. 109, p. 140–148, Apr 2003.

[72] J.-W. Chu, B. L. Trout, and B. R. Brooks, "A super-linear minimization scheme for the nudged elastic band method," *Journal of Chemical Physics*, vol. 119, no. 24, p. 12708, 2003.

[73] Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. a. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. a. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. a. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. a. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. a. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. a. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma,

D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. a. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. a. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. a. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. a. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill, and M. Head-Gordon, "Advances in molecular quantum chemistry contained in the Q-Chem 4 program package," *Molecular Physics*, vol. 113, pp. 184–215, 2015.

[74] S. Grimme, "Semiempirical GGA-type density functional constructed with a long-range dispersion correction," *Journal of Computational Physics*, vol. 27, no. 15, pp. 1787–1799, 2006.

[75] R. A. Kendall, T. H. Dunning, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions," *Journal of Chemical Physics*, vol. 96, no. 9, p. 6796, 1992.

[76] D. Breck, *Zeolite Molecular Sieves: Structure, Chemistry and Use*. New York: John Wiley and Sons, 1974.

[77] T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," *Journal of Chemical Physics*, vol. 90, no. 2, p. 1007, 1989.

[78] S. F. Boys and F. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors," *Molecular Physics*, vol. 19, pp. 553–566, 1970.

[79] J. Jiang and S. I. Sandler, "Separation of CO_2 and N_2 by Adsorption in C_{168} Schwarzite: A Combination of Quantum Mechanics and Molecular Simulation Study," *Journal of the American Chemical Society*, vol. 127, p. 11989–11997, Aug 2005.

[80] B. Vujić and A. P. Lyubartsev, "Transferable force-field for modelling of CO_2 , N_2 , O_2 and Ar in all silica and Na^+ exchanged zeolites," *Modelling and Simulation in Materials Science and Engineering*, vol. 24, p. 045002, Mar 2016.

Appendices

A LAMMPS sample input file

```
# ----------------- Init Section -----------------
    units           real
    atom_style      full
    bond_style      hybrid harmonic
    angle_style     hybrid harmonic
    dihedral_style  hybrid fourier
    improper_style  hybrid cvff
    pair_style      hybrid lj/charmm/coul/long 9.0 10.0 10.0
    kspace_style    pppm 0.0001
    pair_modify     mix arithmetic
    special_bonds   amber
# ----------------- Atom Definition Section -----------------

LAMMPS Description

     32  atoms
     30  bonds
     53  angles
     65  dihedrals
     3  impropers

     71  atom types
     807  bond types
     4248  angle types
     641  dihedral types
     38  improper types

# Simulation box from -20.0 20.0 units in all directions
  -20.0 20.0 xlo xhi
```

```
  -20.0 20.0 ylo yhi
  -20.0 20.0 zlo zhi

Masses # AtomType Mass

1 12.01   # Sp2 C carbonyl group
2 12.01   # Sp C
.
.
.
Atoms # NAtom NMolecule AtomType charge x y z

1 1 4 -0.091100 -4.356 2.171 -2.837
2 1 4 -0.114400 -3.561 1.102 -2.109
.
.
.
Bonds # NBond BondType Atom1 Atom2

1 128 24 27
2 128 17 23
.
.
.

Angles # NAngle AngleType Atom1 Atom2 Atom3

1 315 1 2 6
2 315 2 6 9
.
.
.

Dihedrals # NDihedral DihedralType Atom1 Atom2 Atom3 Atom4

1 10 29 21 28 31
2 10 24 21 28 31
```

```
.
.
.

Impropers # NImproper ImproperType Atom1 Atom2 Atom3 Atom4

1 2 30 21 28 31
2 2 21 28 31 30

# ----------------- Settings Section -----------------
# Starting with non-bonded parameters; AtomType1 AtomType2 Pot epsilon sigma
pair_coeff 1 1 lj/charmm/coul/long 0.0860 3.39966950842   # OPLS
pair_coeff 2 2 lj/charmm/coul/long 0.2100 3.39966950842   # cp C DLM 11/2007
.
.
.
# Bond parameters; BondType BondPot ForceConstant EqDist
bond_coeff 1 harmonic 553.0 0.9572   # TIP3P_Water 1
bond_coeff 2 harmonic 553.0 1.5136   # TIP3P_Water 1
.
.
.
# Angle parameters: AngType AngPotential ForceConst EqAng
angle_coeff 1 harmonic 100.00 104.52   # AMBER 1 TIP3P_water
angle_coeff 2 harmonic 0.00 127.74   # AMBER 1
.
.
.
# Dihedral parameters: DihedType DihedPot nTerms ForceConst Multi PhaseAng
dihedral_coeff 1 fourier 1 0.3 2 180.0    #
dihedral_coeff 2 fourier 1 0.0 2 180.0    #
.
.
.
# Improper parameters: ImpType ImpPot ForceConst cos(PhaseAng) Multi
improper_coeff 1 cvff 1.1 -1 2   # JCC,7,(1986),230
improper_coeff 2 cvff 10.5 -1 2   # JCC,7,(1986),230
```

A LAMMPS sample input file

```
# ----------------- Run Section -----------------

thermo 500
thermo_style custom step evdwl ecoul ebond eangle edihed eimp emol etotal
dump 1 all xyz 500 min1_gaff.xyz
min_style sd
minimize 1.0e-7 1.0e-9 1000 10000
```

B CHARMM sample input file

```
 ----------------------Main File----------------------
* Structure Optimization
* Needs external variable "resi"
*
* Requires mol2crd script, molecule structure in mol2 format
* ,parameter files and.str file of the molecule in the same folder
* .str .mol2 file need the same name (resi.mol2, resi.str)
* resi is defined in .str file after keyword RESI
ioformat extended

read rtf card name top_all36_cgenff.rtf
read para card flex name par_all36_cgenff.prm
stream @resi.str

read sequence @resi 1
bomlev -1 ! for 3-membered rings
generate @resi first none last none setup warn
bomlev 0



write coor card name @resi_temp.crd
system "‘echo ./mol2crd @RESI_temp.crd @RESI.mol2 @RESI_init.crd -
| awk ’{print tolower($0)}’‘"
read coor card name @resi_init.crd
write coor pdb name @resi_init.pdb

COOR ORIEnt
```

```
!minimize structure
MINI SD nstep 20000 nprint 20 inbfrq -1 cutnb 14.

!save minimized geometry
write coor card name @resi_min.crd
write coor pdb name @resi_min.pdb

----------------------Stream File----------------------
* Toppar stream file generated by
* CHARMM General Force Field (CGenFF) program version 1.0.0
* For use with CGenFF version 3.0.1
*

read rtf card append
36 1

! "penalty" is the highest penalty score of the associated parameters.
! Penalties lower than 10 indicate the analogy is fair; penalties between 10
! and 50 mean some basic validation is recommended; penalties higher than
! 50 indicate poor analogy and mandate extensive validation/optimization.

RESI  min1  0.000  !  param  penalty=  4.900  ;  charge  penalty=  5.177
GROUP            ! CHARGE   CH_PENALTY
ATOM C1     CG331  -0.269 !    0.300
ATOM C2     CG321  -0.180 !    4.440
.
.
.

BOND H7   O1
BOND H6   C3
.
.
.

END
```

```
read param card flex append

BONDS
CG2D1  CG311   365.00     1.5020 ! min1.o , from CG2D1 CG321, penalty= 4

ANGLES
CG2D2  CG2D1  CG311    48.00    126.00 ! penalty= 0.6
CG311  CG2D1  HGA4     40.00    116.00 ! penalty= 0.6
.
.
.

DIHEDRALS
CG311  CG2D1  CG2D2  HGA5       5.2000  2   180.00 ! penalty= 0.6
CG2D2  CG2D1  CG311  CG321      0.5000  1   180.00 ! penalty= 4.9
.
.
.

IMPROPERS

END
RETURN
```

B.1 MMFF94 file format in CHARMM

```
         All entries in a Merck-format (.mrk) file have the format:
             Line #     # of Lines      Use

               1            1           Header_1
               2            1           Header_2
               3            1           Number of atoms and bonds
               4            n           Data on n atoms
              4+n           k           Bonding data on the 5*k bonds
                                        in the structure  (Each line
                                        contains data on five bonds)
Header_1:
        The format for the first header line is:
                (A70,A10)
        Each field contains the following information:
                column  Description of use
                 1-70    User defined title
                71-72    Present Year (YY)
                73-75    Present Date (DDD)
                76-79    Time of Day (HHMM), e.g., "1709" for 5:09 pm
                   80    Must be a "1" for the file to be valid
Header_2:
        The second header line has the following format:
                (A4,A8,X,A1,X,A65)
        Each of the fields has the following information:
                column  Description of use
                 1- 4    The string "MOL "
                 5-12    User name
                   14    Source of file : (e.g., E for MOLEDIT, C
                         for Cambridge, D for Distance Geometry etc.)
                16-80    Column used by other programs such as the
                         Cambridge Programs and OPTIMOL
Number_of_atoms_and_bonds:
        The format for this record is:
                (I5,X,I5)
        Each of the fields has the following information:
```

```
                column  Description of use
                 1-5     NATOM
                 7-11    NBND
Data_on_atom_n:
        The format for the atom records is:
        (3(F10.4,1X),I5,1X,I2,1X,I1,1X,I5,1X,3A4,F8.4,6X,A4)
        Each of the fields has the following information:
        Columns     Field               Description
          1-10       X                  X coordinate of the atom
         12-21       Y                  Y coordinate of the atom
         23-32       Z                  Z coordinate of the atom
         34-38       Atomic Number      (I5) field containing the type
                                         of atom. (i.e. -- 6 for Carbon;
                                         8 for Oxygen; etc...) A value
                                         of 0 indicates a lone pair.
         40-41        Atom Subtype      (I2) field: on output, contains the
                                         MMFF atom type; is not read on input
            43        Charge Code       Formal charge code of the atom.
         45-49       Sequence Number    (I5) field containing the unique
                                         number by which every atom in
                                         the structure can be identified.
                                         Note: in the CHARMM implementation,
                                         these quantities are not actually
                                         read.  However, the atoms are
                                         expected to be numbered consecutively
                                         from 1 to NATOM and to correspond to
                                         the numbers used in the bond_data
                                         records defined below.
         51-54        Atom Name         Left justified (A4) field.
                                         Should be unique inside a
                                         given residue. (Examples -- "C24 ",
                                         "NH  ", etc...).
         55-58        Residue Name      Right justified (A4) field.
                                         (Examples -- " 123", "123A",
                                         etc...).
         59-62        Residue Type      Left justified (A4) field.
                                         (Examples -- "TRP ", "LYS ",
```

		etc...).
63-70	Partial Charge	(F8.4) field containing the partial charge of an atom in proton units. Note: this entry is written on output, but is not read on input.
77-80	Segment ID	Left justified (A4) field containing a one to four character segment ID identifier.

Note: if any of the A4 fields specified above are blank, the file reader will construct a default name.

Charge_code:

The valid charge codes are:

Code	Charge Code
0	Neutral
1	+1
2	-1
3	Radical
4	+2
5	-2
6	+3
7	-3
8	+4
9	-4

Bond_data:

The block of data at the end of the .mrk file contains the bonding information. Each line of bond data can contain a maximum of five bond definitions. The format for the bond data is:

5(I5,X,I5,X,I2,2X)

For each bond definition,

Field	Description
IFROM	(I5) Sequence number of the starting atom of the bond
ITO	(I5) Sequence number of the terminating atom of the bond
ITYPE	(I2) Order of the bond. (i.e. 1 for a single bond, 2 for a double bond, etc.)

```
                        Bond orders are always integral
Example
min1                                                                    1624617091
MOL SAMUELFR E
   33    31
   -1.3572    -1.2278    -3.2275     6 12 0     1 C1  TESTTEST   0.0000
   -0.1661    -0.4471    -2.6991     6 12 0     2 C2  TESTTEST   0.0000
   -2.2516    -0.5996    -3.2720     1 12 0     3 H1  TESTTEST   0.0000
   -1.1624    -1.6110    -4.2352     1 12 0     4 H2  TESTTEST   0.0000
   -1.5877    -2.0834    -2.5873     1 12 0     5 H3  TESTTEST   0.0000
    0.2176     0.7336    -3.5781     6 12 0     6 C3  TESTTEST   0.0000
   -0.3784    -0.0743    -1.6909     1 12 0     7 H4  TESTTEST   0.0000
    0.7072    -1.1050    -2.6165     1 12 0     8 H5  TESTTEST   0.0000
    1.4996     1.4137    -3.0909     6 12 0     9 C4  TESTTEST   0.0000
    1.7768     2.2553    -3.7300     1 12 0    10 H6  TESTTEST   0.0000
    1.3692     1.7846    -2.0694     1 12 0    11 H7  TESTTEST   0.0000
    2.3254     0.6973    -3.0894     1 12 0    12 H8  TESTTEST   0.0000
   -0.8667     1.7768    -3.6039     6 12 0    13 C5  TESTTEST   0.0000
    0.3725     0.3984    -4.6113     1 12 0    14 H9  TESTTEST   0.0000
   -1.6380     1.9944    -2.6887     8 12 0    15 O1  TESTTEST   0.0000
   -0.8498     2.4886    -4.7399     8 12 0    16 O2  TESTTEST   0.0000
   -1.4778     3.2304    -4.6487     1 12 0    17 H10 TESTTEST   0.0000
    0.0361     4.7851    -1.6814     6 12 0    18 C6  TESTTEST   0.0000
   -1.1810     5.6594    -1.4265     6 12 0    19 C7  TESTTEST   0.0000
    0.0556     3.9293    -0.9982     1 12 0    20 H11 TESTTEST   0.0000
    0.0409     4.3975    -2.7025     1 12 0    21 H12 TESTTEST   0.0000
    0.9609     5.3504    -1.5369     1 12 0    22 H13 TESTTEST   0.0000
   -2.5031     4.9111    -1.5019     6 12 0    23 C8  TESTTEST   0.0000
   -1.2046     6.4866    -2.1458     1 12 0    24 H14 TESTTEST   0.0000
   -1.1062     6.1145    -0.4315     1 12 0    25 H15 TESTTEST   0.0000
   -3.6827     5.7469    -0.9979     6 12 0    26 C9  TESTTEST   0.0000
   -4.6330     5.2136    -1.0972     1 12 0    27 H16 TESTTEST   0.0000
   -3.5463     5.9945     0.0577     1 12 0    28 H17 TESTTEST   0.0000
   -3.7586     6.6837    -1.5592     1 12 0    29 H18 TESTTEST   0.0000
   -2.8370     4.4335    -2.8816     6 12 0    30 C10 TESTTEST   0.0000
   -2.4511     3.9951    -0.8933     1 12 0    31 H19 TESTTEST   0.0000
   -2.2551     4.7715    -3.8952     8 12 0    32 O3  TESTTEST   0.0000
```

```
-3.7129     3.7541    -2.9394     1 12 0    33 H20 TESTTEST   0.0000
16    17  1     16    13  1     14     6  1      4     1  1     32    30  2
10     9  1     13     6  1     13    15  2      6     9  1      6     2  1
 3     1  1      1     2  1      1     5  1      9    12  1      9    11  1
33    30  1     30    23  1     21    18  1      2     8  1      2     7  1
24    19  1     18    22  1     18    19  1     18    20  1     29    26  1
23    19  1     23    26  1     23    31  1     19    25  1     27    26  1
26    28  1
```

C Q-Chem sample input file

```
$rem
MEM_STATIC                      4000
MAX_SCF_CYCLES                  500
SYMMETRY                        false
METHOD                          B97M-V
THRESH_RCA_SWITCH               7
BASIS                           aug-cc-pVDZ
INCDFT                          0
MAX_RCA_CYCLES                  10
SCF_CONVERGENCE                 8
SCF_ALGORITHM                   diis
THRESH                          14
MEM_TOTAL                       16000
MOLDEN_FORMAT                   true
INCFOCK                         0
XC_GRID                         000075000302
SYM_IGNORE                      true
JOBTYPE                         opt
SCF_GUESS                       core
$end


$molecule
0 1
C         -4.01068        0.44394       -2.73325
C         -2.95433        0.60845       -1.64877
H         -4.84253        1.13829       -2.57820
H         -4.41143       -0.57457       -2.71712
H         -3.59288        0.62312       -3.72832
```

C Q-Chem sample input file

```
C         -2.39789         2.03357        -1.59073
H         -2.14357        -0.10194        -1.85126
H         -3.39324         0.34177        -0.68041
C         -1.38838         2.19466        -0.46250
H         -1.84464         1.99175         0.51107
H         -0.98500         3.21296        -0.45442
H         -0.53584         1.51989        -0.59759
O         -1.75951         2.35915        -2.82616
H         -3.21520         2.74868        -1.44743
H         -1.02701         1.71014        -2.94372
C         -1.92885        -1.22224        -6.92553
C         -0.72585        -0.37371        -7.31828
H         -2.81540        -0.93414        -7.49938
H         -2.16058        -1.12261        -5.86073
H         -1.72488        -2.27885        -7.12634
C         -0.97141         1.12814        -7.11432
H          0.14678        -0.70341        -6.74125
H         -0.50000        -0.56571        -8.37436
C          0.17239         1.95280        -7.69695
H         -0.00539         3.02419        -7.55282
H          1.12514         1.70991        -7.21319
H          0.27829         1.76995        -8.77124
C         -1.10687         1.46807        -5.64649
H         -1.90482         1.40038        -7.62304
O         -0.49551         0.95542        -4.72038
O         -1.99595         2.46128        -5.43789
H         -2.01753         2.60533        -4.45537
$end
```